1999

Fuzzy detection of microcalcifications in mammograms

Agustinus Praba Drijarkara
University of Wollongong

Fuzzy Detection of

Microcalcifications in Mammograms

A thesis submitted in fulfilment of the requirement for the award of the degree of

Master of Engineering (Honours)

from

University of Wollongong

By

Agustinus Praba Drijarkara

B.E. (Hons), University of Wollongong, 1995

School of Electrical, Computer and Telecommunication Engineering

1999

Declaration

The work presented in this thesis was carried out by the author in the School of Electrical, Computer and Telecommunication Engineering, University of Wollongong, Australia and has not been submitted to any other university or institute.

Wollongong, August 1999

Agustinus Praba Drijarkara

Acknowledgement

Ad Maiorem Dei Gloriam,

Firstly I would like to express my gratitude to my supervisors, Dr. Golshah Naghdy and Associate Professor Fazel Naghdy, for their guidance throughout this research work, and particularly in the preparation of this thesis.

My study at the University of Wollongong has been made possible by the scholarship from the Indonesian Government, under the STAID-II scheme administered by the Agency for Assessment and Application of Technology (BPPT).

Acknowledgements are also made to Dr. Mary Rickards and her staff from Westmead Hospital, Sydney, Breast Screening Centre, for her assistance.

Thanks to the staff of the School of Electrical, Computer and Telecommunications Engineering at the University of Wollongong, without whose technical and administrative support my research could not have been completed.

Abstract

Breast cancer is one of the most common forms of cancer found in women. Mammography is a method commonly used for detection of breast cancer. A mammogram is a very high spatial resolution X-ray of breast. The mammograms need to be screened for abnormal and possibly dangerous lesions. Computer-aided diagnosis has been an active area of research to detect abnormalities in a mammogram automatically. The focus of this thesis is on automatic detection of microcalcifications in a mammogram using fuzzy image processing. Microcalcification is one of the earliest signs of breast cancer, which is sometimes hard to detect due to its small size, low contrast and blurred boundary. The fuzzy algorithms developed in this work analyse an image at pixel level, detect the abnormalities and identify the edges of abnormalities using fuzzy operators. The developed fuzzy system is applied to a set of high-resolution mammograms in order to validate its performance. The results clearly demonstrate the feasibility and effectiveness of the proposed approach.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Problem Statement

Breast cancer is reported to be the most common form of cancer found in women and also the leading cause of death among other non-preventable cancers [1]. According to recent statistics, one woman in eight in US and one woman in ten in Europe will develop a breast cancer in her lifetime [2]. In Australia, the figure is reported to be about one in 14 women [3]. Although the mortality rate is high, the disease is curable if detected in the early stages.

Mammography is a method commonly used for detection of breast cancer. A mammogram is a very high spatial resolution X-ray of breast. The mammograms need to be screened for abnormal and possibly dangerous lesions. In most of the developed countries, women over 40 are advised to have mammograms once every two years as a precautionary procedure. This is increased to every year after the age of 50. This generates a large amount of images, which needs to be accurately examined and processed.

The effectiveness of such process relies on the ability of the radiologist to identify any existing abnormality. It has been reported that in some studies 20% of

women with breast cancer had negative mammography finding (false negative) [4]. In fact, extensive studies indicate that the radiologists do not detect all the cancer-related information present in a mammogram [5, 6, 7]. The very subtle nature of the radiographic effects are often the source of missed diagnoses, though the role of human error due to varying decision rules, subjectivity of the process, or pure oversight cannot be ignored [8].

Such problems can be overcome if the radiologist is assisted in the screening process by a reliable tool, drawing the attention to more subtle but important features of a mammogram. Such a prospect has encouraged a number of research groups to study the possibility of a computer-based system that can automatically diagnose abnormalities in a mammogram with more consistency or reproducibility. The work reported in this thesis is an attempt in this direction.

In an automated mammogram interpretation system, abnormalities within a densely captured and digitised X-ray image are searched for. Extremely high spatial and grey scale quantisation resolutions are required for digitised mammograms (35 microns and 16 bits). A digitised mammogram could have up to 12 million pixels, quantised to 16 bits of grey level. This represents 24 mega bytes of information to be analysed for each mammogram. This extensive processing requirement differentiates digital mammogram analysis from any other medical image processing tasks. The biggest challenge here lies in the ability to distinguish between noise, normal tissues, and abnormalities such as tumours or microcalcifications.

The X-ray mammograms can reveal many types of breast lesions. Three major types indicating possible breast cancer are:

- microcalcifications
- circumscribed lesions
- stellate lesions

The majority of the lesions detected, however, are of benign nature. Generally, distinguishing benign lesions from malignant ones is the most challenging task in the mammogram diagnosis.

The focus of this thesis is on detection of microcalcifications using fuzzy image processing.

1.2 Fuzzy Image Processing

The majority of the microcalcification-detection methods reported in the literature approach the problem by using a variety of filters to enhance the signal and to suppress the background textures in the original image [9, 10, 11]. In such task, the aim is to isolate useful signals from the background information. The result of the initial filtering stages is an image containing small bright objects on relatively homogeneous background. Subsequent stages attempt to separate the signal from the noise using different methods. The most popular approaches employed include local adaptive thresholding and morphological erosion. In this process, at each stage, the image is modified and the result is passed to the next stage as input. The strength of this approach is that each stage requires only simple algorithms, which can be executed fast.

The primary drawback of such method is the possible loss of the signal representing the microcalcifications due to insufficient enhancement, which results in a false negative diagnosis. Conversely, the background noise can be presented as an important signal due to excessive enhancement, leading to a false positive diagnosis [12]. In addition, in successive enhancements, errors introduced at one stage will be carried forward to the next and possibly amplified. Another problem that needs to be addressed is that a decision is developed through various stages. Once a weak but valid signal is eliminated at one stage, it might be missed altogether in the final decision. This is known as false negative diagnosis and has been recognised as the most dangerous scenario in screening of mammograms.

To remedy such shortcomings, some of the latest work reported in the literature employ fuzzy methods to segment the mammograms and to identify various abnormalities. Fuzzy operators are used to overcome some of idiosyncrasies of the mammogram data.

The term fuzzy set theory or fuzzy logic was first formalised by Zadeh in 1965 [14] and refers to modes of reasoning, which is approximate rather than exact [13]. In fact, human reasoning is usually approximate in nature. To quote from Zadeh, "in fuzzy logic, everything is a matter of degree" [13]. In fuzzy, truth is not measured as mere true or false, but rather as "to what extent it is true".

Fuzzy methods have a great potential in medical applications, as medical reasoning and diagnosis are often fuzzy and uncertain. Particularly, in medical image processing, the objects produced by the imaging systems contain some degree of ambiguity; geometrically, topologically and qualitatively.

Fuzziness also characterises the medical diagnosis processes. Signals and objects that are encountered in medical applications rarely have crisply defined borders. Also, from the practitioners' point of view, their diagnoses are often non-exact, *e.g.*, "it *may* be a tumour" or "this tumour is *rather* benign". That is why the fuzzy set theory has a great potential to model many medical processes.

1.3 Focus of the Thesis

An effective and efficient automated mammogram analysis system has a great potential benefit, due to the ever-increasing number of people participating in breast screening programs and the limited number of trained radiologists available. To this date, automated detection systems of mammogram have not been applied on a large scale and have not been available commercially.

This thesis is focussed on investigating the appropriateness of fuzzy logic in the analysis of digitised mammograms. More specifically, the target is defined as the small lesions in mammogram known as microcalcifications, which is regarded as the main signal of tumour development. The fuzzy logic will be applied to the detection process at the lowest level of the image, namely, at the pixel level. The detection will be carried out by evaluating the structure and intensities of pixels around a suspect pixel and comparing it with those of the known lesions.

While the work reported in this thesis may not be sufficient to develop a system which is ready for clinical validation, it is an attempt to lay a foundation for

future research on this topic by developing an understanding of the potentials of fuzzy set theory in digital image processing.

1.4 Organisation of Thesis

An overview of breast cancer will be presented in Chapter 2. This will include the significance of early detection in breast cancer and the extent of research carried out worldwide to reduce the mortality rate caused by breast cancer. In addition, the target of the study (the microcalcifications) and the specific problems associated with them will be described. A review of studies aiming at designing an automated tumour detection system will be also presented.

Chapter 3 will address the issue of uncertainty in medical imaging and the appropriateness of fuzzy set theory for such applications. The algorithms developed in this work to identify microcalcifications in a mammogram will be described in Chapter 4.

To validate the developed system, a series of experiments were carried out. The experimental procedures and the obtained results are described in Chapter 5. The developed system will be also compared against a crisp method reported in the literature for the detection of microcalcifications.

Finally, in Chapter 6 some conclusions will be drawn and recommendations for future research will be given.

Chapter 2: Background Study

2.1 Introduction

A great deal of research has been conducted regarding breast cancer. This chapter will explain the importance of such research, which is due to the significance of breast cancer in terms of incidence rate and mortality. One way of reducing the mortality rate of breast cancer is to make the diagnosis process more effective and efficient. Towards this aim computer-aided diagnosis systems have been developed. A review of a number of studies carried out in this area will be presented in this chapter.

2.2 General Overview of Breast Cancer

Breast cancer is the most common form of cancer found in Australian women, and also the most common cause of cancer-related death among women. According to the National Breast Cancer Council of Australia [15], from 1990 to 1992, 7516 women on average were diagnosed with breast cancer annually, while 2458 women on average died of the disease each year. The mortality rate between 1982 and 1992 was around 19 to 20 per 100000 woman-years. Breast cancer

accounts for 25% of all cancers in women, causing 18.7% of all cancer deaths in Australia [3].

The cause of breast cancer is still unknown, but there are several factors which are known to increase the risk of developing breast cancer. These include age, hereditary factor, reproductive factors, and possibly dietary factors [16]. Geographically, incidence rates are higher in Northern America, United Kingdom, Northern Europe, and Australia; and lower in Southern and Eastern Europe, Asia and South America [16]. Incidence rates among migrants in Australia follow a similar pattern as the rates in their countries of origin.

There are a number of methods for treating breast cancer, including surgery, radiotherapy, chemotherapy and hormone therapy [17]. However, the success rate of the treatment depends upon the stage at which the cancer is detected. If the cancer is still localised in the breast, the five-year survival rate is around 90%. This figure drops to around 18% if the cancer has spread to other parts of the body. This highlights the importance of early detection of breast cancer.

The most effective method to date for early detection of breast cancer is mammography, due to its ability to detect tumours long before they can be felt. This is done by taking the X-ray image of the breast in two directions: *cranio-caudal* (top-down view) and *medio-lateral* or oblique (side view). In the examination process, the film screens of the mammograms are examined visually by radiologists.

There are several types of mammographic abnormalities which can indicate the presence of breast tumour which can be classified into three groups [8]:

- discrete abnormalities (including calcifications and masses),
- diffuse spatial changes,
- physical changes which occur over time.

Calcifications and masses are the primary signs of malignant carcinomas [18]. However, calcifications, or more specifically clustered microcalcifications, have received more attention due to two reasons:

- their significance as an early indicator of breast tumour, and
- being relatively more difficult to detect compared to the other types of lesions.

Microcalcifications are small deposits of calcium which are formed in the breast tissue. It is believed to be "the result of active secretory process by tumour cells, rather than mineralisation of necrotic tissue" [19]. They are considered to be the main indicator of breast cancer because a significant proportion (30 to 40%) of carcinomas are found to have microcalcifications, which lead to the detection of these carcinomas in early stages [20].

Although microcalcifications can signal the presence of tumour, not all of them are malignant. Some of the visible microcalcifications are benign. Malignant calcifications can be distinguished from benign ones by their shape, size, and

distribution pattern. Benign microcalcifications tend to be larger, rounder, fewer, and less variable in size [21]. On the contrary, malignant microcalcifications usually are numerous and vary in shape and size [18].

Microcalcifications become clinically significant only when they appear in clusters of three or more within an area of 50×50 mm. The probability of malignancy is even higher if the cluster contains branching and casting microcalcifications. Another indication of malignancy is when the presence of microcalcifications is not associated with a mass [18].

The occurrence of carcinomatous microcalcifications in mammograms, which appear as "fine grains of salt", was first reported in 1950 [18]. They appear in mammograms as small spots which are brighter than the background, due to their higher radio-opacity than the surrounding tissue.

Sometimes microcalcifications are difficult to see because of their small size and low contrast. The average diameter is only about 300μm, and most of them are smaller than 700μm [9]. Those of diagnostic importance are generally smaller than 500μm [11]. Some microcalcifications have low contrast either because they are located in the denser parts of the breast, or because they are still in the early stage of their development. Another reason for the low contrast is the decrement in the radiation dose of the mammographic images [22].

The large number of mammograms required to be examined is another difficulty associated with mammography. The majority of mammographic images are normal. That is, only a small proportion of mammograms in a screening might

contain microcalcifications. For example, in the British screening program, less than 1.5% of women in the screening program exhibited abnormalities which are significant enough to be followed up by a biopsy [8]. The screening process has become an overwhelming task for the limited number of radiologists available. Moreover, research suggests that the sensitivity of the radiologists diminishes with an increase in the number of images to be examined [23]. This reduces the effectiveness of the screening process.

In the developed countries with high rates of breast cancer incidence (for example, the USA, the UK and Australia), their governments or health authorities have been promoting regular breast screening for every woman in the high-risk age group. A single baseline mammogram is suggested between ages 35 and 40, followed by regular mammography every two years. Women over 50 should undergo routine screening every year. As the participation rate in such programs increases, the number of mammograms will also increase significantly. Consequently, there is a need for a more efficient method for examining the mammograms.

Mammogram analysis is very labour intensive because two radiologists are often required to read a mammogram to reduce the risk of misdiagnosis. Since only a small number of mammograms contains abnormalities, radiologist mostly examine normal images. An automated system can be designed to examine the images and draw the attention of the experts to the suspicious regions, allowing them to concentrate on these suspicious cases.

In recent years, computer-aided diagnosis systems have been introduced to assist in the mammogram analysis process, in order to reduce the cost and maintain its effectiveness. The ultimate aim is to develop an automated system which can perform a thorough examination of a mammogram for any kind of abnormalities. The intention is not to replace the human expert altogether; rather, the automated system will act as an aid or assistant to the expert. Such system can also provide the expert with a second opinion [24, 22]. This is partly because of the complexity and the subjective nature of the diagnosis process. To date, none of the reported methods can claim to have a 100% success rate.

2.3 Image-Processing Approaches to Mammogram Analysis

Image processing is defined in the literature (*eg.* [25, 26]) as a method to rearrange information in a digitised image for two purposes:

1. Improving the visual appearance for human perception;

2. Preparing the image for measurement and analysis in machine perception.

The first purpose is aimed at enhancing the more useful information and suppressing the less useful parts, so that the important features can be perceived better. The analysis and interpretation of the pictorial information are done by the human observer. The second purpose, on the other hand, is aimed at utilising the computer to perform the analysis and interpretation automatically. An example is the detection of the presence of an object of interest, or measuring the features of the object or structures.

Many mammogram-analysis schemes have been designed with the second purpose in mind, *i.e.*, to automatically detect and report the presence of microcalcifications, and to verify the malignancy of the microcalcifications. Even though the research so far has not produced a fully reliable system, the developed automated methods can provide a 'second opinion' to the diagnosis of the radiologist [22, 24]. In this way, they can also act as a 'prompter'; that is, to draw the attention of the radiologist to more suspicious regions in an image. The final decision, however, rests with the radiologist.

Analysis of digitised mammogram can be carried out through three stages [2]:

1. detection and segmentation of objects present in an image;

2. extraction of the features of the objects;

3. classification or assessment of the malignancy of the objects.

The first stage is aimed at locating the lesions or abnormalities in the image which resemble a microcalcification. However, not all lesions are necessarily a cause for concern. Many of these lesions are perhaps just normal or benign changes in the breast tissue, or even dust or dirt which are embedded on the film during mammography. In the later stages of the analysis, such objects are classified according to their potential malignancy.

There are certain problems associated with microcalcifications which are not usually encountered in other medical image techniques [9]. One of those problems is the variability of the normal breast tissue making up the background of a

mammogram. Some microcalcifications may be located in a denser part of the breast than others, in which case absolute intensity of both the object and the background will be higher, while the contrast will be lower. The inhomogeneous intensity of the background does not allow a global thresholding based on the pixel intensity. Another problem is the small size of the objects of interest, which requires sufficiently high spatial resolution. The low contrast between some of the objects and the background is another problem which requires high grey-level resolution.

Considering the above problems, a good microcalcification detection method should satisfy the following requirements [9]:

- Being insensitive to variation of intensity in the background. This means that the absolute intensity of either the background or the objects is less important than the difference or contrast between the two.

- Being adaptive to noise level within a neighbourhood. The presence of noise should not reduce the sensitivity of the algorithm.

- Being robust relative to the size and shape of abnormalities. Microcalcifications appear in various shapes and sizes. Hence the algorithm should not be trained to find objects with a particular shape or an exact size. Although, the search can be limited to a range of values reflecting the dimensions of the clinically important objects.

An extensive review of mammogram analysis schemes has been carried out by Astley [8]. This study also includes methods developed for non-

microcalcification abnormalities such as well-defined tumour or masses [27, 28]; comparison between successive mammograms [29]; and comparison between the right and left breasts [30]. The following sections will review other methods which are reported in the literature, especially those which are specifically aimed at detecting microcalcifications.

2.4 Microcalcification detection methods

The microcalcification detection methods reported in the literature can be divided into two main groups: those that search for the microcalcifications by removing the background texture and/or enhancing the small objects, and those that perform segmentation on the image and extract the features of the segmented region to find the microcalcifications.

2.4.1 Background removal and object enhancement

In this approach, the detection is performed in two stages. The background structures are removed from the image and small objects resembling microcalcifications are enhanced. Then the extracted objects are further processed and enhanced, to remove false objects and noise.

This approach uses the signal-processing paradigm, in which the components of the image are defined in terms of signals with different frequencies. An image or an image object is said to have a high frequency if the pixel intensity changes rapidly in the spatial domain. Low-frequency images, on the contrary, are made of

pixels whose values differ only slightly. In the case of microcalcification detection, the object of interest, which is the microcalcification itself, is considered to be a high-frequency signal. On the other hand, the background, which is made up of the normal breast tissue, is considered to be a low-frequency signal.

In signal processing, it is common to use filters to separate the components of a signal which have different frequencies. The simplest forms of filters are the high- and the low-pass filters. As the name suggests, a high-pass filter preserves the high frequency components while suppressing the low frequency ones. The opposite applies for the low-pass filters. Image processing, as an extension of signal processing, also uses filters in the spatial domain to separate the components of an image. There are various theories and methods to implement frequency-based filters in the spatial domain.

2.4.1.1 High-frequency analysis

The background removal process is essentially a high-pass filtering process. In the spatial domain, there are various techniques and methods to implement filtering.

A technique called the *difference-image* is used by Nishikawa [10], which uses both high-pass and low-pass filters. The raw image is processed separately through each filter. The low-pass filter suppresses small objects, while the high-pass filter emphasises them. The low-pass filtered image is then subtracted from the high-pass filtered image, leaving the objects on a relatively plain background. For

the high-pass filter, a fixed 3x3 kernel or filtering mask is used. The weakness of this method is that lesions larger than the kernel size will diminish both in size and intensity.

The difference-image technique is also used by Dengler [9] for separating the objects. For the high-pass filter, a Gaussian-based filter is used. This has an advantage over Nishikawa's kernel since the width of the filter is adjustable to suit objects with different sizes. Still, in each operation only one particular size of the filter can be used. If the sizes of the objects vary widely, objects much larger or much smaller than the filter size might be missed or significantly diminished.

Another variation of the difference-image such as used by Mascio [31] uses two filters for analysing the high-frequency signals: a round *high-emphasis* filter, which essentially is the same as the difference-image technique, and *texture gist* filter, which combines the erosion and dilation processes. The first filter emphasises objects with rather crisp boundaries larger than several pixels, while the second emphasises small and textured details in the image.

High-pass filtering can also be done using wavelet transform, as done by Lo [7]. A brief description of wavelet is presented in the next section. It is reported that a 3-level wavelet transform is performed, after which the lowest frequency compartment is removed and subsequently the image is reconstructed with inverse wavelet transform. The high-pass filtered image is then thresholded at several levels to extract all the suspected spots. The use of thresholding method has a drawback which will be discussed later.

2.4.1.2 Texture analysis

Another method of distinguishing microcalcifications from the background is the analysis of texture. Texture can be defined as a unique variation in brightness [25]. One way to analyse texture is by considering the local non-uniformity of the pixels, as done by Cheng [32]. It is postulated that microcalcification pixels can be distinguished from normal breast tissue pixels according to their local non-uniformities, which are calculated from their local variances (a statistical quantity of the neighbouring pixels' grey level values). In Cheng's method, the pixels with low variances are removed, leaving some curvilinear background structures, which is further removed by calculating the length and elongation of the curves. The report does not explain how these quantities are determined.

Wavelet analysis is another, more effective method for analysing textures and has emerged lately as one of the popular methods for analysing mammograms [11, 22, 24, 33]. Full explanation of the wavelet theory is beyond the scope of this thesis. For the purpose of background, it suffices to say that any signal has two characteristic components: frequency and time. In image processing, time domain is analogous to spatial domain. The wavelet transform is able to extract the frequency and time components from a particular signal, using a basis function which has two parameters: resolution (or scale) and translation [22]. The challenge is to find the correct values for these parameters in order to emphasise the objects whose parameters resemble those of a microcalcification.

Nesbitt [11] uses wavelet transform in a similar manner as Lo [7] which is discussed previously. The wavelet transform is performed at several levels, then certain scales are emphasised with different weight factors and the image reconstructed via inverse wavelet transform. The segmentation of the objects is done by adaptive thresholding, based on the mean and standard deviation of pixel intensities in the sub-image that contains the object.

The work reported by Yoshida [24] is an extension of the work reported in [10], discussed above. Basically the approach is the same as in their previous work, except that the wavelet transform replaces the high-pass filter. A supervised learning method is used to tune the wavelet to get the correct parameters. The wavelet transform method is reported to have a sensitivity of 95%, compared to 85% produced previously.

Naghdy *et al.* [33] use two-tier Gabor wavelet and neural network environment to detect the microcalcifications. The Gabor wavelet is not only sensitive to frequency but also to the orientations of signals in the spatial domain. The parameters of the Gabor wavelet are tuned using a fuzzy adaptive-resonance-theory neural-network classifier. The reported classification rate is about 93%.

Dhawan [19] combines wavelet decomposition with second-order grey-level histogram statistics to analyse textures. The first is used to represent the local texture of the microcalcification area, while the latter is used to represent the global texture.

Another approach to texture analysis is fractal theory [34]. Using this method, the background can be identified as the regions with high local self-similarity (compared to the microcalcifications which have less structure) and therefore can be removed. The result is reported to be equal or better than wavelet method. It is also reported that this method can remove more background structures than wavelet, but is not as good as wavelet in preserving the overall shapes of the spots.

Textural features can also be computed from the co-occurrence matrices, which are a measure of how often pairs of grey-levels of pixels, separated by a certain distance along a certain direction, occur in the image. This is reported in [35]. The co-occurrence matrices are computed from the output signals of a quadrature mirror filter (QMF) bank, which consists of low and high-pass filters.

2.4.2 Post-processing of background-removed images

Background removal processes such as described in the previous sub-section do not yet produce the end result. Although most of the irrelevant background has been removed, some 'ghosts' or residue of the background textures still remain. Also, there may be some other objects which resemble the real lesions and therefore pass through the high-pass filters, but are not of interest to diagnosis. These might be objects (such as dirt or dust) which emerge during the filming process, grains on the film screen, or electrostatic noise produced in the digitisation process. They are usually either very small or have high contrast relative to the background. Post-processing is required to remove the irrelevant objects.

Most of the post-processing methods use morphological operators to remove noise based on a specific size; for example, objects smaller than three pixels. The contrast of the objects relative to the background is another selection criteria.

Binary thresholding is almost always used at some stage of the post-processing. Because of this, the end result is binarised to zero and one, representing the background and the objects, respectively. It has been argued [35, 19] that binary segmentation is not suitable for mammogram images due to poor contrast in some objects, which makes it difficult to accurately draw the border between the object and the background. Binarisation has one disadvantage: it removes the inherent uncertainties of the objects. For example, some objects have diffuse boundaries. From a medical point of view, the crispiness (or the fuzziness) of an object's boundary may have some significance in distinguishing benign from malignant lesions.

Another disadvantage of microcalcification detection through background removal approach is that the image is altered at each stage of the process. At each stage some information may be lost. This information, which contributes to the uncertainties of the object, may be of secondary significance; nevertheless, it may help to create a more accurate classification.

2.4.3 Contour analysis

An approach which does not involve background removal is proposed by Bankman [12]. In order to explain this approach, the mammogram is considered as

a topographic map where the height of the "landscape" is represented by the intensity values of the pixels. In this visualisation, microcalcifications will appear as hills or peaks protruding from a relatively flat background.

The "hills" can be detected by firstly drawing lines connecting adjacent pixels with the same intensities, commonly known as contours. The area around a peak will have several concentric contours. A microcalcification can be identified from the contours by evaluating the number of concentric contours and the size of the outermost contour.

To determine whether a set of concentric contours is a microcalcification, three features are evaluated: departure, prominence and steepness. Departure is a measure of the sharpness of the perceived edge of the object, which also marks the object's perimeter. Prominence reflects the relative brightness of the object compared to the background. Steepness is the gradient of the landscape, which gives a measure of whether the object has a sharp edge or a diffused one.

This approach has the potential to extract the information about the image, in its original state.

Another approach with similar strategy is used by Cairns [36], which also uses contours. Unlike Bankman's method where the contours are derived from the intensity of the pixels, in Cairn's approach the contours are generated by connecting contour-cues, which are derived from the gradients of the pixels. The gradients themselves are extracted with Sobel edge detection method. The

drawback of this method is when the edge of the object is not clearly defined, in which case there are several possible contours for one object.

2.5 Object classification

The objective of analysing microcalcifications is to classify the detected objects into one of the following categories [20]:

- highly suspicious for malignancy,
- definitely benign, or
- indeterminate.

In the reviewed microcalcification detection methods, the classification is usually based on the features extracted from the objects after the detection stage. Some of the reviewed work in this area, however, focuses solely on the classification methods without describing the detection, segmentation and feature extraction processes.

2.5.1 Features used for classification

The features used for classification can be divided into two groups [37]. The first consists of features with direct correlation with the characteristic radiographic signs of malignancy known to radiologists. This means that the features can be described by the expert, at least qualitatively. These include, for example, the

number of microcalcifications in a cluster, the size and shape of individual microcalcification, and so on.

The second group of features are those which give good separation between distributions of malignant and benign cases. These are the features which may not be readily obvious or could not be perceived consciously by a radiologist. These include, for example, features which have to be calculated from the statistics of the numerical values of the pixels in the microcalcification area. The two parameters of the wavelet transform [24, 33] can be considered in this group.

The features from the first group are easier to define, but may be difficult to measure. For example, measurement of object size requires the outlining of the objects, which may be difficult for those with fuzzy boundary. On the contrary, the features from the second group may be easier to compute, but are not readily obvious as to whether they provide the effective discriminating criteria. The reported detection methods use the features either from the first group only, or a combination of the first and second group features. Cairns [36] uses seven features: size, shape, brightness, homogeneity, edges, and clustering; all of which fall into the first group because they are readily observable visually.

A work by Aghdasi [38] has evaluated over 100 features from individual and clustered microcalcifications, including: photometric variables (*e.g.*, mean and variance of intensities, histogram parameters, *etc.*); size variables (area, perimeter, and radius); shape variables (including compactness and elongation); roughness variables; and some other. Using the classification method that is employed in this

work, five features emerge as having the highest discriminating power. These are: number of microcalcifications, minimum inertia, minimum compactness, normalised standard deviation of area and normalised standard deviation of perimeter.

Hall [1] uses seven core features including area, shape, average edge strength, edge strength variation, contrast, object standard deviation plus background standard deviation, and the Laws energy features.

Because different microcalcification-detection methods use different sets of features as the classification criteria, it is not clear which features are the most effective for identifying microcalcifications. Each method uses different techniques to measure the same features. For example, measuring the size of an object from its binarised representation and from the fuzzified representation may give different results. It seems that the choice of discriminant features depends upon the way the features are extracted and the classification methods used.

2.5.2 *Classification methods*

Classification can be done simply by determining whether the values of the features fall in the corresponding acceptance range, as done by Bankman [12]. The acceptance range is usually determined from the statistics of a number of samples or a training set. Bankman [12] reports that with this simple approach and using three features, the algorithm is able to detect the clusters and reject the other structures and artefacts with no false cluster. However, the test set consists only of

two images. It is not certain how the system will perform with a larger number of test cases.

A clustering scheme called *Isaac* is proposed by Estevez [39]. This scheme classifies candidate microcalcifications according to five textural features and four difference histogram features. The clustering process comprises of selective clustering and interactive adaptation. In the clustering stage, the data is separated according to the feature space into groups representing the true and the false objects. The interactive adaptation stage allows the radiologist to improve the clustering result by identifying false objects and excluding them from the true object clusters.

In Aghdasi's work [38] which is mentioned earlier, there were initially over 100 features to consider. These features are computed statistically using a commercial statistical analysis package to calculate the Fisher statistics, which indicate the discrimination power of each feature. Using this method, the feature with the highest discrimination power has a 89.7% classification accuracy.

A follow-up of Bankman's work is reported in [40], in which the classification is done by a feedforward neural network. The features, which are used as inputs for the neural network, are the same as in the previous work, plus four additional features: distinctness, compactness, mean slope and area of the base contour. The modified system was tested on 18 images. The result was a 93% sensitivity with 1.56 false cluster per image.

Neural network is also used by Jiang [37] in conjunction with the detection method similar to that reported in [10]. The inputs of the neural network are the following features: cluster area and circularity, number of microcalcifications per cluster, and per unit area, mean distance between microcalcifications, mean area and effective volume of microcalcifications, and the second highest irregularity measured for microcalcifications in a cluster. The system correctly classifies 38 out of 40 malignant clusters, and 34 out of 67 benign clusters.

Fuzzy logic and its variations have also been used, *e.g.*, [2, 41, 42]. Murshed *et al.* [41, 42] uses a fuzzy-ARTMAP based classification system for detecting cancerous cells in microscopic images. Although the object is different, the system probably can be modified for classifying mammographic images as well. The fuzzy ARTMAP system is actually a derivative of the neural network and not a 'real' fuzzy logic.

A fuzzified version of a well studied decision tree classifier, C4.5, is used in Hall's work [1] to classify microcalcifications based on the seven core features described earlier. The C4.5 decision tree is originally a binary tree system. In this work, the output of the decision tree is fuzzified to improve the classification accuracy.

Bothorel *et al.* [2] is one of the first to implement a 'real' fuzzy segmentation and classification algorithm for analysing microcalcifications. The fuzzy segmentation method preserves the ambiguities of an object by extracting several possible contours which mark the border of the object. Each of the possible

contours will have different membership value, denoting the probability of the contour of being the true border. Therefore, in the classification stage, the classification of one object is based not only on the features of a single contour, but also of all the possible contours.

A neurofuzzy system for modelling input-output data has been investigated by Bridgett [43]. This system is intended to form part of an intelligent oncology workstation, which not only analyses the lesions but can also suggest possible treatment methods.

2.6 Conclusion

This chapter has reviewed significant research work automatic detection of microcalcifications. There are many approaches which have been investigated, each with varying degree of effectiveness. It is not possible at this stage to accurately compare the effectiveness of each method, because the employed validation methods are different. In addition, the test images have different degrees of subtlety and resolutions. An algorithm may perform differently if tested using images with different spatial or intensity resolutions.

Detection methods which require image enhancement (which cause the alteration of images) may lose some information from the original image. Other detection methods operating directly on an unaltered image may have a better chance of detecting the objects correctly.

Chapter 3: Uncertainty in Medical Images

3.1 Introduction

Human-based decision making and diagnosis in natural sciences such as medicine is usually prone to a degree of uncertainty. Mathematical modelling of such processes using precise logic and mathematics will not be successful as the approach does not match the nature of the process. Fuzzy logic on the other hand provides the means to describe and manipulate uncertainty. This has made the fuzzy set theory an attractive method to model diagnosis process in medicine including the analysis and interpretation of medical images.

This chapter will explore the nature of uncertainty in medicine in general and in medical image processing in particular. It will then review various methods used to model uncertainty with more emphasis on fuzzy set theory.

3.2 Uncertainty in Medical Analysis

The process of diagnosis and decision making in medicine are inherently affected by uncertainty and imprecision. Uncertainty and imprecision comes from two sources [44]:

1. the uncertainty in the medical knowledge about the relationship between symptoms and disease, and

2. the uncertainty in the knowledge about the object under observation (in this case, the patient or a substance taken from the patient).

The process of diagnosis begins with the perception of the symptoms, followed by recognition of the symptoms and conclusion about the disease. This assumes that the information about the disease and symptoms is already in the reference knowledge. The diagnosis process, in which the observed knowledge is compared with the reference knowledge, is subjected to the following elements of uncertainty [45]:

- imprecise information
- inaccurate information
- missing information
- conflicting information.

Imprecise information is the result of subjective and qualitative perception of the objects in the real world by human. For example, "a peach" can be described as a kind of fruit with a rather round shape, a slightly pointed end and a large seed. In such description the features are all expressed in qualitative terms (*round, large*) which are imprecise. Furthermore, additional modifier words are used to describe the intensities of the qualitative terms (*rather* round, *slightly* pointed), which add to the uncertainty.

In the above example, the description about an object (a peach) is vague, due to the inherent fuzziness in the human perception. The description is not specific because the same description can be applied to a different but similar kind of fruit (for example, a plum or an apricot). The information provided by the description is not enough to distinguish a peach from a plum or an apricot (a result of incomplete or inaccurate information). For this purpose additional information will be required; for example, by adding statements about other features in the description (*e.g.*, colour), or by describing a feature more precisely or as a comparison (*e.g.*, a plum is *smaller* than a peach).

Recognition is basically a process of comparing the description of a new object with the descriptions of known objects (objects which have been encountered previously). Uncertainty in the recognition process comes as a result of comparing the unknown object with similar, rather than identical, objects [46]. In a normal situation, two objects may be considered identical if they have a high degree of similarity between them. In reality, though, natural objects are never exactly identical.

Uncertainty in the decision making results from conflicting information received from several experts. Quite often different experts express different opinions about a situation.

3.3 Imprecision in Medical Images

Medical diagnosis carried out based on examination of medical images can suffer from the inherent imprecision of the images. The sources of imprecision in medical images are the blurred boundaries of objects, inter-individual variation, vague description, reduction of resolution, and complexity of the shapes.

3.3.1 Blurred Boundaries

Objects in a medical image often have blurred or fuzzy boundaries. This makes it difficult to define the objects' physical dimensions. One of the most common image-processing tasks in medical applications revolves around defining the boundaries between different objects or regions in an image. The fuzziness of the boundary is the result of several causes:

- blurred boundary between tissues of different organs,
- blurred boundary due to movement during the imaging process,
- scattering of X ray by dense tissues,

- noise produced during the digitisation process (in the case of digitised images).

The boundaries between microcalcifications and the background tissue in mammograms, in particular, quite often are fuzzy. Determining the boundary of microcalcifications is a crucial task as the classification process assumes the presence of a border. The shape and size of a microcalcification can only be determined after the border is defined.

3.3.2 Inter-individual Variation

Since every individual is unique, the same organs from different people are never exactly the same, although they have similar features. The shape of a hand, for example, is the same for everyone. But the length of the digits, the width of the palm, and the pattern of the lines on the palm have a wide variation from one person to another. Inter-individual variation also applies to lesions.

3.3.3 Vague Description

As explained in the previous section, objects are often described with vague and inexact descriptions. Vague description of an object is caused by the inter-individual variation. This is where medical image analysis differs from other areas of image processing. For example, in the analysis of an image from a production line as part of a quality control procedure, the slightest deviation in an object from the prescribed template can immediately be identified as an error. In medical

analysis, on the other hand, some tolerance must be allowed to accommodate the variation.

3.3.4 Reduction of Resolution

In the transformation from analogue into digital images, some information may be lost due to changing of analogue or continuous signal into discrete picture elements or pixels. The size of the imaging element determines the size and spatial resolution of the pixels. A small object may be under-represented if the spatial resolution is too low. Another aspect of resolution is the chromatic and luminance resolution, or the number of colours and brightness intensities that can be represented in the image. This is determined by the sensitivity of the imaging equipment and the number of bits representing each pixel.

Resolution is particularly relevant to microcalcification detection in mammograms because of their small size and low contrast.

3.3.5 Complex Shapes

The shape of an organ or lesion may be too complex to be exactly described. For example, imaging the brain stem is a major problem. Fortunately, microcalcifications have a simple shape for digital imaging and do not cause any problem in this respect.

3.4 Handling Uncertainties in Knowledge Base

In an automated medical analysis process, the relationships between symptoms and disease (the knowledge base) must be expressed symbolically in a way that it can be manipulated and interpreted digitally by a computer. Since these relationships can rarely be modelled as mathematical equations, the knowledge base usually takes the form of "production rules" such as the following:

$$\text{IF } X = A \text{ THEN } Y = B$$

where X represents the symptoms and Y represents the disease. The rule implies that the conclusion (Y = B) is true if the antecedent or condition (X = A) is satisfied. The rule also contains uncertain components in the antecedent, conclusion and inference [45]. As an example, consider the following rule:

If (body temperature is hot) then (fever is present)

Since temperature is normally measured numerically (*e.g.*, temperature = 37°C), the antecedent of the above rule cannot be evaluated without first defining what is "hot". In conventional logic system (or "crisp" logic), this is done by defining a limit to differentiate different situations. For example, suppose that any temperature higher than 37°C is considered hot. The rule then becomes:

If (body temperature > 37°C) then (fever is present)

The uncertainty in the antecedent comes from the fact that a little difference in temperatures (*e.g.*, between 36.9 and 37.1°C) may not have any significant effect

as far as the disease is concerned, yet the inference of the rule would produce totally different results. The temperature 36.9°C will be considered cold, while 37.1°C is hot. Different persons may have different physical endurance level. Thus an exact limit in the antecedent cannot always be defined for the whole population. Another component of uncertainty comes from the inaccurate measurement of the temperature, which is subject to the precision of the measuring equipment.

The conclusion of the crisp rule often does not provide a very informative result. A grading of the conclusion could be more helpful in deciding the treatment. For example, the presence of fever can be expressed in different degrees: no fever, mild fever, high fever, *etc.*

The inference of the rule also contains a degree of uncertainty. Is fever always present when body temperature reach 37°C? A high temperature might be a result of another physical condition, so the inference is not always true. An indicator of the probability of the inference being true is certainly required. According to Leung [46], there are some common approaches to dealing with uncertainty in the knowledge base: Bayesian approach, certainty factors, Dempster-Shafer theory of evidence, and fuzzy logic. Two of the most commonly used are explained below.

3.4.1 Certainty Factor

Certainty Factor (CF) is used in many expert systems where uncertainty in the reasoning process is recognised [45]. For example, suppose that the rule in the

example above is defined with a CF of 0.9. However, it does not provide a clear indication whether the CF indicates the degree or intensity of the situation stated as the conclusion (*i.e.*, how bad is the fever?), or whether it reflects the probability of the inference being true if the antecedent is satisfied. This system is used, for example, in the MYCIN decision support system.

3.4.2 Fuzzy Set Theory and Fuzzy Logic

Fuzzy logic handles imprecision in the inference process by handling the input and output not as precise numbers, but rather as degrees of truth. The fuzzy set theory makes it possible to estimate imprecise information as fuzzy values which can be used to compute the fuzzy expected value, leading to a crisp decision [47].

Consider again the above example. Crisp logic approaches the concept of "hot" by defining a crisp border to distinguish "hot" and "not hot" (*i.e.*, cold). Fuzzy logic, however, can approach this problem by assigning different degrees of truth for the statement "temperature is hot". That is, each temperature has a *membership value* in the *set of hot temperatures*. The membership value for each temperature can be determined by a *membership function*. Figure 3.1 illustrates a possible membership function to define hot temperature. According to this membership function, for example, 36.9°C has a membership value of 0.9 in the set of hot temperatures. This means that the statement "36.9°C is hot" is true to a degree of 0.9, and therefore 36.9°C is not totally cold.

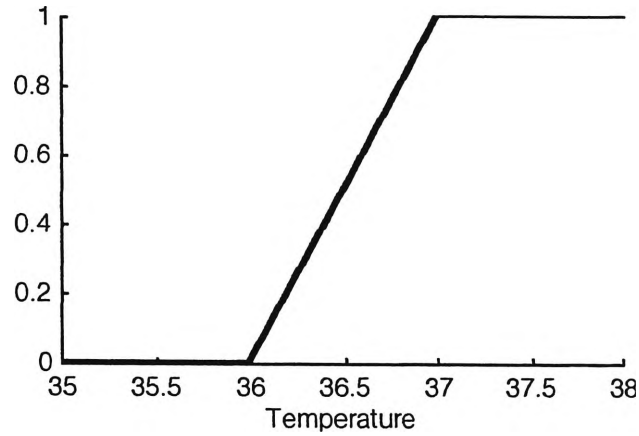

Figure 3.1. Membership function for "hot temperature"

By using membership values, the uncertainty in the input (or the antecedent) of the inference rule can be reflected in the output of the inference rule. By defining membership functions to the fuzzy concepts that make up the knowledge base, an expert system that follows closely the human thinking and reasoning method can be developed.

3.5 Applications of Fuzzy Set Theory

The early development of fuzzy logic had been primarily in control systems. A wide range of commercial products have been designed with fuzzy logic as their controllers; from toasters to washing machines; from lift controllers to an electric subway train in Japan [48]. Fuzzy logic became so popular to the general public particularly in the early 1990's, that the name fuzzy logic was widely used (and misused) in the advertisement of home appliances to make the products more appealing. Such products, compared to the "conventional" ones, were claimed to be more versatile, more effective and more efficient.

Several studies have been conducted to evaluate the usefulness of fuzzy logic in medical-related applications. Early studies were focussed on medical decision support systems (MDSS) in diagnosing symptoms. In another field of engineering, fuzzy logic has also been applied for image processing and analysis. As the two areas of study converge, fuzzy set theory has also been applied in medical image analysis.

One important reason for investigating the use of fuzzy logic in image processing is stated by Tizhoosh: "... fuzzy logic provides us with a mathematical framework for representation and processing of the expert knowledge" [49]. Thus fuzzy logic provides the means to bridge the gap between the imprecision of the linguistic concepts and the numerical nature of the digitised images.

3.5.1 Implementations of Fuzzy Set Theory

There are several applications of fuzzy set theory in image processing and understanding. Some of the techniques, in order of theoretical and practical relevance to this work, includes fuzzy clustering, rule-based approach, fuzzy geometry, measure of fuzziness, fuzzy measure theory, fuzzy morphology, and fuzzy grammars [49]. The first two, being the most studied and investigated, are explained below.

3.5.1.1 Fuzzy clustering

Clustering is the process of separating a number of data into several groups, so that the information in the same group have similar properties. There are two types of clustering, based on the initial assumption: *unsupervised clustering*, where there is no initial knowledge or assumption about the number of clusters and the defining criteria; and *supervised clustering*, where there is some form of intervention to arrive at an expected outcome. Figure 3.2 illustrates clustering of a set of data, represented in two-dimensional space as two parameters x_1 and x_2. The same set of data can be clustered according to two similarity measures; angular displacement in polar coordinates, and Euclidean distance in Cartesian coordinate [50].

Clustering can also be divided into *hard* or *crisp* clustering and *soft* or *fuzzy* clustering. In crisp clustering, each data point belongs to one cluster only. However, there are some situations in which a data point is located roughly halfway between two clusters. Such data may belong either to one of the clusters, or to none at all. Assuming the data has to belong to a cluster, crisp logic usually resolves this by assigning the data to the closest cluster. Fuzzy logic, on the other hand, acknowledges the ambiguity of the data and allows the data to have partial membership to any cluster.

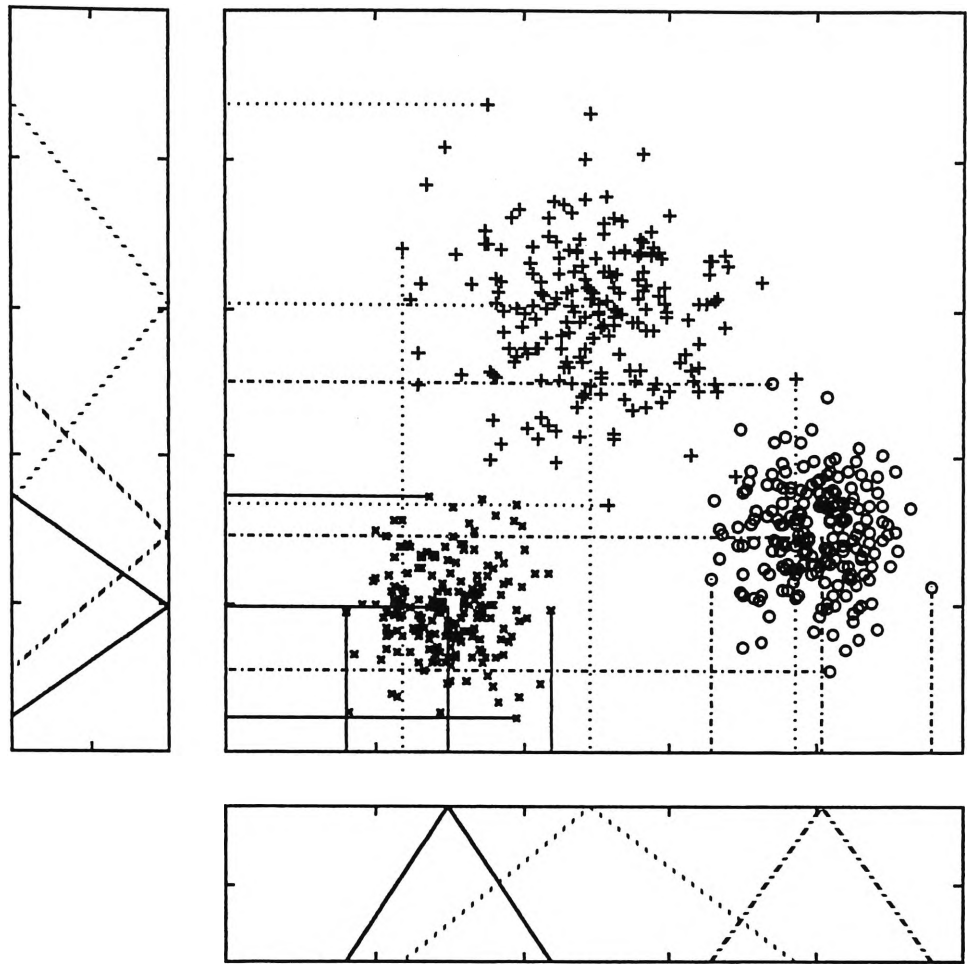

Figure 3.2. Fuzzy clustering of three classes of data. Each data point has two variables, which are represented by the axes. The fuzzy membership function of each class of data in each of the variables can be derived from the distribution of the data in the respective axes.

Examples of fuzzy clustering implementation are summarised below:

- Segmentation of image [51]: Segmentation is the first step in many image analysis applications which require the image to be segmented into different regions, where each region represents a unique object or area. This can be done by clustering the pixels based on grey-scale, colour, location, *etc.*

- Grouping several elements of image which make up a bigger object [52].

- Object recognition [53]: This is carried out by measuring certain dimensions of the object in question and comparing them to those of the template to find a similarity.

3.5.1.2 Rule-based system

Rule-based systems are the common form of expert system. As explained in Section 3.4, a fuzzy rule-based system consists of a set of *if-then* rules which are translated from the expert knowledge. This is different from fuzzy clustering which does not necessarily require *a priori* knowledge. However, the concept of fuzzy clustering can be incorporated in building a rule-based system to define the membership functions.

Rule-based systems are commonly applied in decision support systems. In image processing, they are often used for object recognition. Although it can be also applied to low-level processes.

3.5.2 Decision Support Systems

Khamseh [54] reviews some of the MDSSs, including MYCIN, Oncoin, Iliad, Meditel, DXplain, QMR; and compares them with a fuzzy-based diagnosis system. The fuzzy-based system operates on a knowledge-base which is a collection of rules acquired from the experience and observations of the expert. The conclusion of this study is that each of the different methods used in MDSSs has its own merits and limitations, but fuzzy-based systems show better overall performance.

Other decision support systems are reported in [44, 55, 46]. A system called System Z-II is reported by Leung [46], which is one of the first expert system able to handle fuzzy concepts expressed in natural language according to the expert knowledge. The system can also easily mix fuzzy and normal terms.

Sidaoui [44] reports of a diagnostic methodology which is based on a fuzzy representation of symptoms, taking into account their relative importance. This system can accommodate the approximate knowledge of multiple experts, even if conflicts of opinion occur.

A fuzzy hierarchical approach to medical diagnosis is reported by Zahan [55]. The advantage of this system is that the diagnosis does not simply say whether a particular disease is present or absent, but more importantly, it gives the possibility degree of the diagnosis. The possibility degrees rank from *impossible* to *extremely possible* or *sure*.

3.5.3 Image Processing and Analysis

Image processing can be separated into two general levels: high and low. High-level processing deals with objects and structures, while low-level processing deals with pixels which make up the image. As a comparison, in the human visual system the high-level processing is the more conscious processes, compared to the low-level which is less conscious.

3.5.3.1 Object Recognition or Classification

Recognition, often used in the context of "image understanding", is based on a comparison of features extracted from regions or segments of an image. It is assumed that the image is already segmented, where each segment represents a single entity in the real world. As recognition is a rather conscious process, fuzzy-based recognition system are more often in the form of rule-based systems. Examples of fuzzy-based object recognition and classification are described in [56, 57, 41].

3.5.3.2 Pixel Neighbourhood Operators

A neighbourhood operator is one which determines the properties of a pixel based on the properties of other pixels in the surrounding area. Using neighbourhood operators, the following low-level processes can be achieved: edge detection [58, 59, 60, 61]; noise elimination [62, 63]; segmentation [64]. A neighbourhood operator is basically a set of rules which describe the patterns or forms that exist among a pixel and its neighbours.

3.5.4 Fuzzy-based Medical Image Analysis

Fuzzy logic has gained enough popularity among medical image analysis researchers, which is evident from the number of reports in this area. Quite a number of work in this area are focussed on MRI images, particularly in the

segmentation of brain images [62, 65, 66, 67]. The common approach to segmentation in these reports is based on clustering the pixels according to the grey-level values. Then an additional rule-based neighbourhood operator is applied to classify the ambiguous pixels.

Other examples of fuzzy logic in medical image analysis include determination of pedigree by extracting edges from bone image [68], chromosome recognition [69]; and non-invasive examination of foetuses [70].

3.6 Summary

The design of an expert system in the area of medical image analysis must take into consideration the fuzziness that is inherent in the images and in the reasoning processes. The underlying theory of fuzzy logic arguably suits this task very well, although there is still a great deal of work to be done to prove this. However, results obtained so far provide a promising future.

Chapter 4: Design of Fuzzy Microcalcification Detector

4.1 Introduction

In this chapter, the design of the fuzzy microcalcification detection system will be explained. Initially an overview of the system will be presented, including the basic concepts and paradigms, and an outline of the detection process. Then the details of the detection processes will be explained.

4.2 Overview of Fuzzy Detection Approach

4.2.1 Conceptual Approach

The methodology developed in this study is based on a number of concepts developed and validated in previous studies. The work is in particular inspired by a topographic approach introduced for visualising the digitised mammogram, and a fuzzy similarity analysis used for comparing objects in a mammogram against a reference model.

In the microcalcification detection scheme proposed by Bankman [12], the digitised mammogram is visualised as a topographic map, where the intensities of the pixels are represented by the height of the landscape. In this perspective the microcalcifications would appear like hills or peaks projecting prominently against the relatively flat landscape of the background tissue. The hills are characterised by several properties: the height, relative to the base; steepness of the slope; and diameter of the outline. Based on this visual concept, a model of the object of interest can be developed. Figure 4.1 shows a typical microcalcification image and its topographic representation.

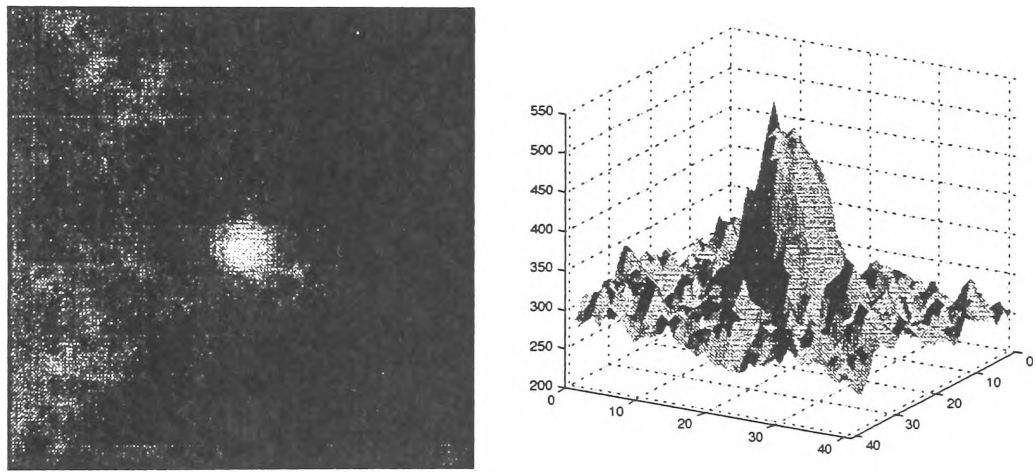

Figure 4.1. An example of a typical microcalcification in a segment of image measuring 100×100 pixels (3.5×3.5mm), and its topographic representation.

This study is focussed at recognising microcalcifications at the pixel level by comparing a model of the microcalcification with the objects in question. The comparison is performed by placing a window of observation around a candidate object and comparing the pixels within that window to the pixels in a template model. Each pixel in the window will be assigned a degree of similarity to the corresponding pixel in the template. The combined degrees of similarity from all

pixels determine the degree of similarity of the object with the model; in other words, the likelihood of the object being a microcalcification. This concept of "similarity analysis" has been used in several fuzzy-based edge detection schemes, in particular [61] and is illustrated in Figure 4.2.

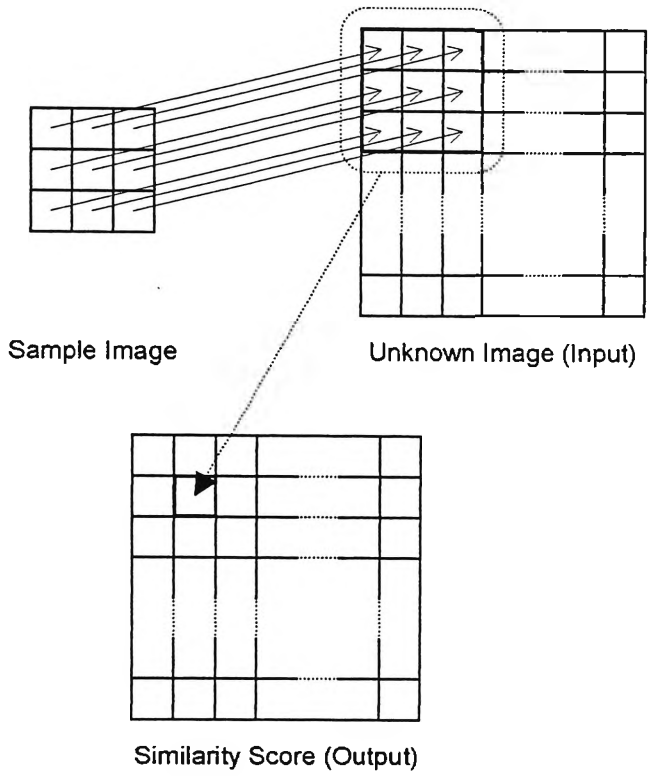

Figure 4.2. Similarity analysis of pixels between an image and a sample. A segment of the image is compared pixel by pixel to a template of the sample, and a similarity score is assigned to the pixel at the centre of the segment.

The fundamental principle of the comparison process in the object recognition is that the similarity is evaluated as a fuzzy concept rather than crisp. This means that the model is also defined in terms of fuzzy sets.

4.2.2 Description of Target

The model of the microcalcifications will be developed by incorporating the features which are generally accepted as the defining characteristics of microcalcifications. In particular, these features are those which are used by radiologists in their description and can be perceived visually. This study does not attempt to assess the useability of unknown features. The features to be used are size, shape, relative brightness, and the presence of border.

4.2.2.1 Size

It is assumed that the size of the microcalcifications will vary within a certain range and therefore the detection will be aimed at objects with a size within this range. This assumption will be held only for the development of the prototype. This will have to be discarded in the real-life application to account for objects with a size outside the range. The limitation due to this assumption and the significance of this limitation will be discussed in the validation of the method.

4.2.2.2 Shape

Although quite a few microcalcifications have a linear shape, most of them can be approximated to be round. The detection system, however, is designed to be insensitive to slight deviations in roundness.

4.2.2.3 Brightness

Microcalcifications appear as bright spots in mammograms, due to their higher opacity compared to the background tissue. In the digitised mammogram, the brightness intensities are represented by grey-scale values, which have a number of levels depending on the resolution of the system.

The absolute brightness, or the grey-scale value alone, is not sufficient to identify the spots. The grey-scale value of the spots and the surrounding pixels may be different from one part of the image to another, depending on several factors including the density of the local tissues, the stage of the microcalcifications development, and the effects of digitisation process. The lesions must be analysed in terms of their relative brightness compared to the local background. This requires normalisation of the grey-scale values of the local region, by scaling the values within the region into the full grey-scale range of the display system.

Most of the reported studies confirm that the brightness within the boundary of a microcalcification has a uniform pattern with no specific texture.

4.2.2.4 Presence of Border

A border is a line which separates a dark region from a bright region (in the case of monochrome image). The line itself, however, does not exist such as in comical drawing; it is merely the visual perception of the transition from dark to

bright (or *vice versa*). It should be noted that the border has a fuzzy nature; although it can be perceived visually, it cannot be defined with absolute precision.

The use of border as a defining feature is seldom mentioned explicitly; however, it has been used directly or otherwise in some studies. In the early discovery of microcalcifications by Leborgne, they are described as "fine grains of salt" [18], which implies that the grains are separable from the background. In Bankman's study [12], the border is measured in terms of the steepness of the slope. Detection schemes based on high-pass filters in principle search for the high-frequency component of the objects; in other words, the borders.

4.2.3 Outline of Detection Process

The detection is carried out in several stages, each using a particular operator. The objective of each stage is to identify which pixels have high probability of being microcalcifications and which pixels do not. In this way, the non-microcalcification pixels are successively eliminated from the processing chain. Being a fuzzy system, the elimination is not performed in the manner of a crisp classification; each detection stage merely assigns a low score to a pixel which does not quite exhibit the properties or characteristics being detected at that stage. The actual elimination of the pixels will be carried out at the end of the chain, according to the score which is assigned by the last stage in the chain.

The main operator in this detection process is the fuzzy peak detector, which is a neighbourhood operator. In common windowed operations, the operator is

usually applied to every pixel in the image by shifting the operator window over the entire image. However, for this application it is not necessary to do so. An initial selection of pixels is performed to select the "local peaks" as candidate pixels for the fuzzy peak detection. The fuzzy peak detector evaluates each candidate pixel and its neighbours to determine if the object represented by the pixel resembles a peak, and assigns a membership function μ_{peak} to the candidate pixel. Candidate pixels with high scores of μ_{peak} are evaluated by a fuzzy edge detector, which determines if the corresponding objects have an observable border. The strength of the border around a candidate pixel is represented by a membership function μ_{edge}, which is the final score for the pixel. A pixel which has a high score for μ_{peak} and also has a high score for μ_{edge} is therefore more likely to be a part of microcalcification.

4.3 Fuzzy Neighbourhood Operator

In image processing operations where the shapes of objects or structures are relevant, a class of operators called *neighbourhood operators* are often used. They can be used to detect the presence of a certain shape, or to change the shape of objects. Hence an alternative name "morphological operator" is also used.

The term neighbourhood refers to the way these operators work by examining the distribution pattern of intensities in a small neighbourhood around a pixel. When the image is binary, the operator can be implemented with a mask or template which has the likeness of the shape of interest. The template can also be

implemented as a virtual mask, which is a description of the shape, specifying which pixel should be black and which should be white. This concept can be extended with fuzzy logic, by defining "black" and "white" as fuzzy sets.

As an example of neighbourhood operators and a basis for developing a peak detector for microcalcification detection, a fuzzy-based edge detection algorithm by Li [60] is described here. The size of the edge operator in this algorithm is 3×3 pixels and the members are labelled as in Figure 4.3. The pixel being evaluated is labelled *Q*.

3	4	5
2	Q	3
1	8	7

Figure 4.3. Window labelling in Li's edge detector.

An edge occurs in an image where the intensity changes abruptly. More specifically, an edge pixel is located between a region containing dark pixels and another containing bright pixels. There are several possible configurations of dark and bright pixels around an edge pixel, as depicted in Figure 4.4. A black box represents a dark pixel, a white box for a bright pixel, and a grey box for a "don't care".

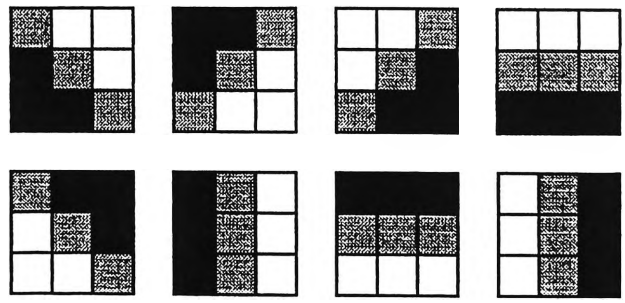

Figure 4.4. Possible pixel configurations around an edge pixel [60].

Since there are eight possible configurations of edge pixel, eight different templates or detection rules will be required. Each rule will be stated in the following form: "if a window has a similar configuration as Template *K*, then the centre pixel is an edge pixel". To describe the process in more detail, the top left template in Figure 4.4 will be used as an example.

To evaluate the similarity of an image window with the template, the following variables are defined:

- *dif_lum(i)*: the intensity difference between pixel *Q* and pixel *i*; $i = 1 \ldots 8$. *dif_lum* is characterised by fuzzy labels *Neg* and *Pos*.

- *Neg* and *Pos*: fuzzy membership values defining the degree of pixel *i* being dark or bright; if *dif_lum(i)* has a large positive value, then the statement "*dif_lum(i) is Pos*" is true to a high degree. The membership functions for *Neg* and *Pos* are shown in Figure 4.6.

- *lum(Q)*: the output of the rule, defining to what degree pixel *Q* is an edge pixel.

- *Black*: a fuzzy membership value defining how dark pixel *Q* should be in the output; an edge pixel is represented by a dark pixel in the output.

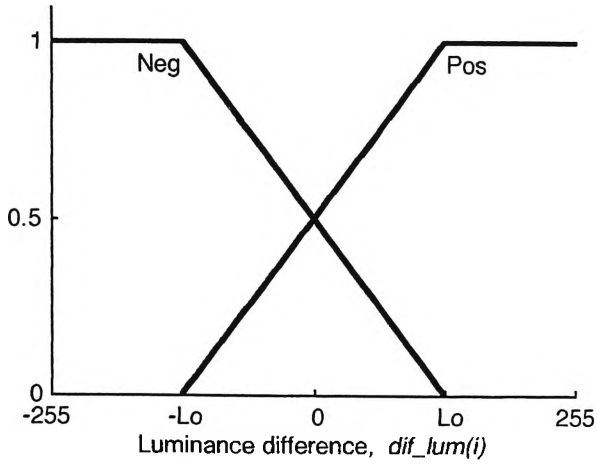

Figure 4.6. Membership functions for *Neg* and *Pos*, as functions of luminance difference *dif_lum(i)*.

The rule for detecting an edge pixel, based on Template 1 (the top left template in Figure 4.4), can then be stated as Rule 4.1 below.

Rule 4.1:

If *dif_lum*(1) is *Neg*
and *dif_lum*(2) is *Neg*
and *dif_lum*(8) is *Neg*
and *dif_lum*(4) is *Pos*
and *dif_lum*(5) is *Pos*
and *dif_lum*(6) is *Pos*
Then *lum(Q)* is *Black*,
Else *lum(Q)* is *White*.

The above rule can be illustrated as in Figure 4.7.

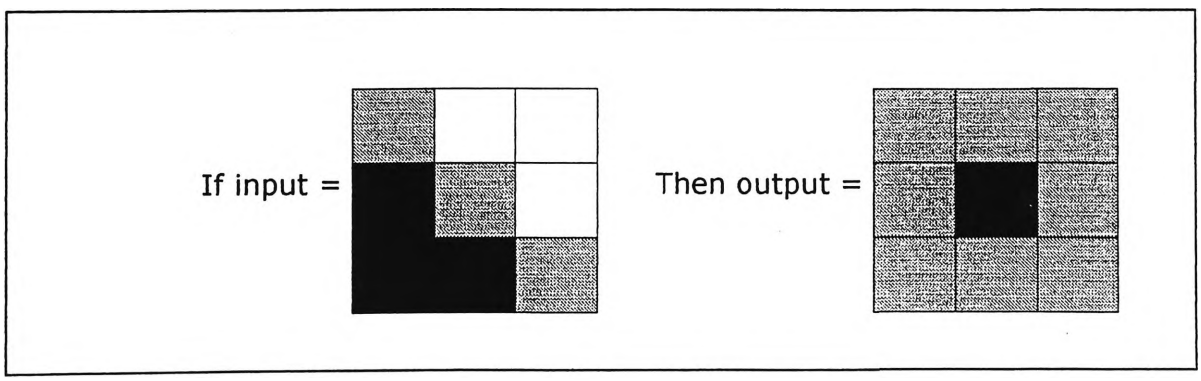

Figure 4.7. First rule for detecting edge pixel.

There will be eight rules similar to Rule 4.1 to describe all the edge templates. Each rule is evaluated by calculating *dif_lum(i)* for each pixel in the corresponding template, then applying the fuzzy membership functions for *Neg* and *Pos* on the appropriate pixels, as specified in the rule. For example, when evaluating Rule 4.1, the following steps are taken:

- calculate *dif_lum(i)* for pixels 1, 2, 4, 5, 6, and 8
- calculate *Neg* for *dif_lum(1)*, *dif_lum(2)* and *dif_lum(8)*
- calculate *Pos* for *dif_lum(4)*, *dif_lum(5)* and *dif_lum(6)*
- evaluate the antecedents, then draw the inference for this rule.

Finally the results of the eight rules will be combined. If the image contains an edge pixel, then one of the rules (never more than one) will yield a high score. Otherwise, if it is not an edge, all the rules will come up with low scores.

In summary, in order to apply a fuzzy neighbourhood operator to perform a certain function, a number of rules must be defined to describe all the possible pixel configurations, and each rule will have several antecedents to describe the individual pixels. The antecedents are evaluated with fuzzy membership functions and the result is drawn through fuzzy inference.

4.4 Local Peak Selection

The proposed microcalcification detector is a neighbourhood operator which is designed to generate a high score when the operator window is applied on the centre pixel of a microcalcification. When applied to other pixels in the microcalcification, the score should be lower, and very low with non-microcalcification pixels. By applying the operator on every pixel in the digitised mammogram, all pixels can be classified into microcalcification and non-microcalcification. However, it may not be necessary to operate on all pixels, if the objective is simply to determine the locations of the microcalcifications. It should be sufficient to apply the operator once only on every prospective microcalcification, on a pixel which has the highest probability of scoring. Such pixels are the brightest ones in the microcalcifications; the peak of the "hill".

The first criterion for candidate pixels is being a local peak which is a pixel brighter than the immediate neighbours. For example, in a 3×3 window, if the centre pixel is brighter than any of the peripheral pixels, then it is a local peak. The peak of a microcalcification is by nature a local peak. By choosing the local peaks as candidate pixels, each microcalcification would be represented by at least one pixel. Pixels on the slope of the "hill" are not local peaks and will not be included in the next stages. This selection criterion ensures that every possible peak is examined by the peak detector while greatly reducing the number of candidate pixels, and reducing the time required to process the image.

4.5 Fuzzy Peak Detector

4.5.1 Preliminary Design

The fuzzy peak detector being proposed is evolved from the edge detector described in Section 4.3. The peak detector needs a bigger operator window so that a whole microcalcification can fit in. Suppose that a typical microcalcification is about 3 pixel wide, as depicted in Figure 4.8. A 9×9 operator window would be sufficient for this hypothetical example. A tentative labelling scheme for such window is illustrated in Figure 4.9.

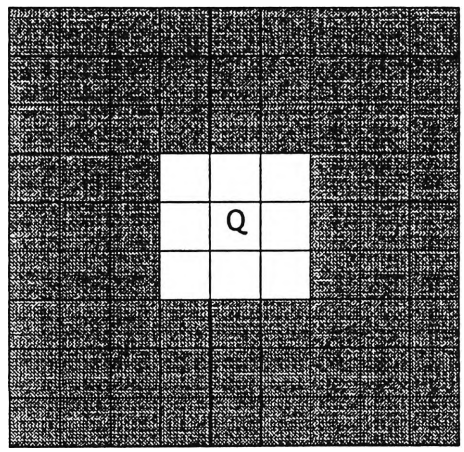

Figure 4.8. A hypothetical microcalcification image.

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}
a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
a_{28}	a_{29}	a_{30}	b_1	b_2	b_3	a_{31}	a_{32}	a_{33}
a_{34}	a_{35}	a_{36}	b_4	Q	b_5	a_{37}	a_{38}	a_{39}
a_{40}	a_{41}	a_{42}	b_6	b_7	b_8	a_{43}	a_{44}	a_{45}
a_{46}	a_{47}	a_{48}	a_{49}	a_{50}	a_{51}	a_{52}	a_{53}	a_{54}
a_{55}	a_{56}	a_{57}	a_{58}	a_{59}	a_{60}	a_{61}	a_{62}	a_{63}
a_{64}	a_{65}	a_{66}	a_{67}	a_{68}	a_{69}	a_{70}	a_{71}	a_{72}

Figure 4.8. Tentative window labelling scheme for the microcalcification detector.

A rule to detect a microcalcification of the shape depicted in Figure 4.7 will be defined as in Rule 4.2. This rule has 80 antecedents in total, in accordance with the number of neighbour pixels in the observation window. Eight antecedents describe the bright pixels (labelled b_i) and 72 other describe the dark pixels (labelled a_i).

Rule 4.2:

If		pixel b_1 is *Bright*
	and	pixel b_2 is *Bright*
	and	...
	and	pixel b_8 is *Bright*
	and	pixel a_1 is *Dark*
	and	pixel a_2 is *Dark*
	and	...
	and	pixel a_{72} is *Dark*,
Then		Q is probably a microcalcification.

As stated previously, if a different pixel configuration is possible, another rule has to be defined. However, the variation in pixel configuration for microcalcifications is very wide. Microcalcifications with different sizes, shapes and orientations will have greatly varying pixel configurations. Furthermore, because the size of the operator can be quite large (since a microcalcification can be as big as 25 pixel wide), each rule will have a large number of antecedents, in

accordance with the number of pixels in the operator. The great number of rules combined with the large number of antecedents in each rule requires a modification on this design.

4.5.2 Modified Design

Consider a window of observation centred around a microcalcification. The pixels within the window will exhibit the following properties:

Definition 4.1:

- pixels which are near the centre, are almost as bright as the centre; and
- pixels which are far from the centre, are much darker.

For simplicity, in the following descriptions the term "bright" will be used to describe the pixels which are *almost as bright as* the centre, and "dark" for those which are *much darker* than the centre.

There are four fuzzy variables used in Definition 4.1 to describe a neighbour pixel: *near*, *far*, *bright*, and *dark*. This leads to four possible combinations of variables: near and bright; far and dark; near and dark; and far and bright.

Definition 4.1 states that any pixel around a microcalcification must be either (near and bright), or (far and dark). Since a pixel cannot be both (near and bright) and (far and dark), these two combinations of variables can be examined at once with an "or" operation. In this light, Rule 4.2 can be redefined as Rule 4.3.

Rule 4.3:

If	pixel *1* is (near and bright) or (far and dark),
and	pixel *2* is (near and bright) or (far and dark),
and	...
and	pixel *N* is (near and bright) or (far and dark),
Then	*Q* is probably a microcalcification.
Remark: N is the number of neighbour pixels in the operator window.	

By redefining the rule as above, only one rule is required to detect microcalcifications with different shapes, sizes or orientations (within a certain limit of variation). The number of antecedents is not reduced, and each antecedent is now four times more complex because four fuzzy variables have to be examined. However, the advantage of having one general rule to define any form of microcalcifications outweighs the cost of overloading the antecedents.

4.5.3 Mathematical Calculation

To explain the mathematical computation of the fuzzy peak detector, the following symbols need to be defined:

- Q: the candidate pixel; that is, the pixel being examined;
- P: the set of neighbour pixels around Q, within a window of observation of size $N \times N$ centred at Q;
- p_i: a member of P; so that $P = \{p_i \mid i = 1, \ldots, M\}$; M = the number of neighbour pixels within the window of observation;
- $int(p)$: intensity of pixel p;
- d_i: spatial distance between pixel p_i and Q;

- g_i: relative grey-level intensity of pixel p_i with respect to Q; $g_i = \frac{int(p_i)}{int(Q)}$;

- $\mu_{NEAR}(d_i)$ and $\mu_{FAR}(d_i)$: membership functions describing the spatial distance of p_i (which could be near or far); and $\mu_{DARK}(g_i)$ and $\mu_{BRIGHT}(g_i)$: membership functions describing the relative intensity of p_i (which could be bright or dark). The basic forms of these membership functions are shown in Figure 4.9.

- $\mu(p_i)$: a membership function denoting the similarity of p_i to the corresponding pixel in the template; in other words, the degree to which p_i satisfies the antecedent of Rule 4.3.

- $\mu_{PEAK}(Q)$: the membership degree of Q in the set of peak objects, denoting how similar the object represented by Q is to a microcalcification.

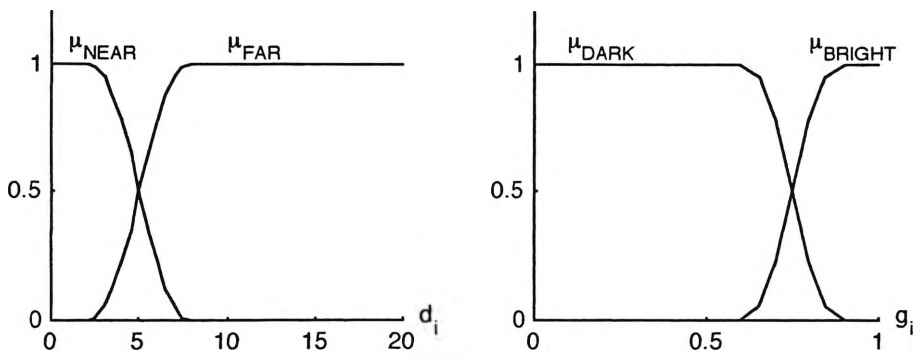

Figure 4.9. Basic forms of membership functions describing *near*, *far*, *bright* and *dark* for the spatial distance and relative intensity of the neighbour pixels.

The fuzzy peak detector operates in succession on each of the candidate pixels determined in the previous step. When evaluating a candidate pixel Q, the operator takes into account a segment of the mammogram centred around this candidate pixel. The neighbour pixels p_i around Q are then compared to the

microcalcification template, as described by Rule 4.3. The similarity $\mu(p_i)$ of a neighbour pixel to the corresponding pixel in the template is determined by the antecedent statement of the rule: "pixel p_i is (near and bright) or (far and dark)", which can be formulated as in the following equation:

$$\mu(p_i) = [\mu_{BRIGHT}(g_i) \wedge \mu_{NEAR}(d_i)] \vee [\mu_{DARK}(g_i) \wedge \mu_{FAR}(d_i)] \quad [4.1]$$

where the symbols 'v' and '∧' denote the fuzzy 'or' and 'and' operators, respectively.

Since each antecedent of Rule 4.3 can be expressed as in Equation 4.1, the whole rule can be expressed as follows:

$$\mu_{PEAK}(Q) = \mu(p_1) \wedge \mu(p_2) \wedge \ldots \wedge \mu(p_M) \quad [4.2]$$

which can be solved by calculating the mean value of $\mu(p_i)$, for all i. The value of $\mu_{PEAK}(Q)$, therefore, denotes the degree of possibility of the candidate pixel Q being part of a microcalcification.

4.6 Fuzzy Edge Detection

A common approach to edge detection is by considering the gradient or the intensity difference between adjacent pixels. An edge occurs where the gradient is relatively high. In many edge detection methods, such as the Sobel operator, the gradient is extracted by means of convolution using particular masks, followed by binary thresholding. In this work, the conventional edge detection method will be

used with a modification to implement the fuzzy principle. The edge detection is performed through three steps: gradient extraction, ridge thinning, and fuzzy thresholding. Finally, the strength of the edge is measured from the extracted edge pixels. The intermediate results of the edge extraction are shown in Figure 4.10.

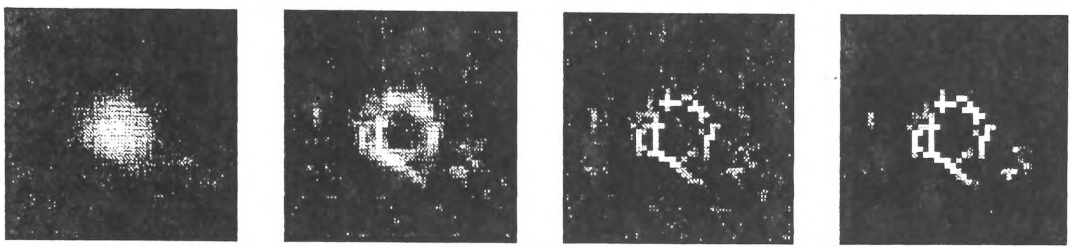

Figure 4.10. Edge detection process: (a) Original image; (b) Gradient of image (a); (c) Thinned gradient; (d) Fuzzy-thresholded gradient, revealing the edges.

4.6.1 Gradient Extraction

The most common gradient extraction method is the Sobel method, which involves convolution of the image by using the convolution masks depicted in Figure 4.11, known as the Sobel operators. Each of these masks in principle calculates the intensity of gradient in one direction. The overall gradient is the square root of the sum of the squares of the two directional gradients. However, in many practical applications the largest directional gradient is considered as the overall gradient [25].

1	2	1
0	0	0
-1	-2	-1

-1	0	1
-2	0	2
-1	0	1

Figure 4.16. Sobel convolution masks.

In this work, four convolution masks are used. In addition to the orthogonal masks, two diagonal masks are used to calculate the gradients in the diagonal directions. The diagonal masks have a factor of $\sqrt{2}$ to compensate for the spatial distance between diagonally adjacent pixels. The use of diagonal masks gives a more accurate gradient representation, compared to only two.

The result of the convolution is shown in Figure 4.14 (b).

0	0	0
1	0	-1
0	0	0

$\frac{1}{\sqrt{2}}$	0	0
0	0	0
0	0	$-\frac{1}{\sqrt{2}}$

0	1	0
0	0	0
0	-1	0

0	0	$\frac{1}{\sqrt{2}}$
0	0	0
$-\frac{1}{\sqrt{2}}$	0	0

Figure 4.17. Convolution masks used in this work.

4.6.2 Ridge Thinning

The result of the gradient extraction is an image with some structures resembling ridges. It can be seen that the edges in the original image coincide with the ridges; and sharper edges correspond to more prominent ridges.

To get a more accurate representation of the edges, the ridges are thinned using a morphological operator. The ridge thinning operator has a 2×2 window which is scanned across the gradient image. The pixel with the smallest value

among the four pixels in the window is zeroed. This leaves the vertices or the skeletons of the ridges, which outline the edges in the original image. This is shown in Figure 4.10 (c).

4.6.3 Fuzzy Thresholding

The ridges in the gradient image are not only produced from the edges of the microcalcification but also from the background variation. It can be seen that the residual ridges of the background are not as bright as those from the actual edge, and can be removed by fuzzy thresholding.

Fuzzy thresholding is a special case of fuzzy clustering and it simply means that there are only two clusters. The threshold level in this case does not refer to a single cut-off point as in the case of crisp binary thresholding. Rather, the threshold level is defined as a range of values where the fuzzy membership function changes from zero to unity. In this case, the two clusters represent the set of edge pixels and the set of non-pixel edges, respectively. The edge pixels are those with *high* gradient *intensity*. The fuzzy threshold level for *high*, however, cannot be defined uniformly for all cases because the gradient intensity varies from case to case. The threshold level (in other words, the membership function) must be defined adaptively.

An *ad hoc* method is used to define the fuzzy threshold level, based on the statistical information of the ridge intensities. Visual observation suggests that the intensity of the edge ridge is significantly higher than the rest.

Figure 4.13 shows a histogram of intensity of the ridges in Figure 4.10 (c). The pixels with the lowest values from the ridge image are removed iteratively, until most of the residual ridges are removed without significantly degrading the edge ridges. This gives the crisp threshold level, which is marked with h in the histogram. The intensities of the edge ridge therefore are higher than h. In the histogram, the edge intensities form the right-hand tail, which is almost separated from the bulk of the distribution.

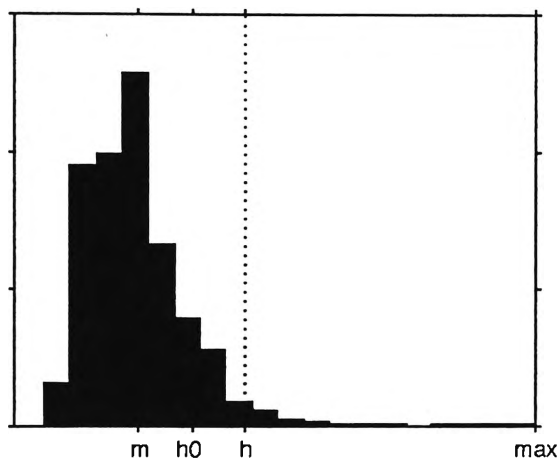

Figure 4.13. Histogram of the ridge intensities from the ridge image in Figure 4.10 (c).

The histogram can be approximated by normal distribution, whose characteristic width is measured by the standard deviation. Therefore the optimum fuzzy threshold levels will be defined in terms of the mean m and the standard deviation σ in the form:

$$h = m + k\,\sigma \qquad [4.3]$$

The coefficient factor k is determined from a number of microcalcification samples which become the training set. The value of h is determined manually while the values of m and σ are computed automatically. This gives a range of values for k. Then using fuzzy clustering method, the threshold levels for k are determined, and subsequently, the threshold level in terms of h is obtained.

4.6.4 Edge Strength Calculation

The final step in determining if an object has an edge is by measuring the edge strength. The strength of the extracted edge can be measured from the completeness of the edge around the object. If the edge is broken in several points, the strength will be lower. Measuring the edge completeness can be done simply by counting the edge pixels, after establishing the normal pixel count for the edge. Pixel count, however, is affected by the size of the object. A large object with half-complete edge may have a higher edge-pixel count than a smaller object with complete edge. To overcome the issue of object size, the pixels are given weighting factors to normalise the strength of each pixel.

The edge-pixel count of an object is approximately proportional to the perimeter (the approximation due to the pixelation effect), which in turn is proportional to the radius. Therefore to normalise the edge strength, the edge pixels are weighted with the inverse of the spatial distance between that pixel and the centre pixel.

4.7 Defuzzification

Typically, a fuzzy logic system would have complementary fuzzification–defuzzification stages, in which information or data from the real world is transformed into fuzzy terms as input for processing and inference, and then the output is transformed back into numerical data that can be directly applied to engine settings, translated into grey-level values of processed images, or as other kinds of numerical input. The two fuzzy detectors in this application basically perform the fuzzification function. However, for this application it is not necessary to defuzzify the output. The reason is that the output is already in the form of scalar value, which can be used directly with the ROC curve method for determining the effectiveness of the system in distinguishing microcalcifications from other objects. Although the original data is in the form of grey-level values, it is not necessary to translate the fuzzy values of the pixels back into grey-level values, because the final judgement is done on individual pixels, not on the basis of the image as a whole. Therefore, in this work a specific defuzzification is not carried out.

4.8 Summary

The design of the fuzzy microcalcification detection system has been explained in this chapter. It consists of several processes which perform successive selection of the possible microcalcification objects. The selection processes are carried out so that the uncertainties of the objects are taken into account throughout the processing chain. The numerical results and the overall performance of the method will be presented in the next chapter.

Chapter 5: Validation

5.1 Introduction

To validate the fuzzy microcalcification detection system, the system is tested on a number of digitised mammograms. This chapter will describe the criteria and methods used for testing the algorithm. A method called ROC curve analysis is used to assess the quality of the detection system. The performance of the fuzzy detection system will be compared to a non-fuzzy microcalcification detection system which is discussed in Chapter 2. Results of the tests will be presented and discussed.

5.2 Performance Measures

A simple measure of detecting quality is 'accuracy', which is the ratio of the number of correctly detected objects to the total number of objects being examined. However, this is not an adequate indicator because it is affected by the prevalence of the targets in the sample population [71]. A more meaningful way to measure the quality of detection is to measure two properties of *sensitivity* and *selectivity* separately.

Sensitivity refers to the ability of the system to detect the actually positive lesions and is measured in terms of true positive (TP) and false negative (FN) rates. A "true positive detection" is when an actual object is correctly identified. A "false negative detection" is when an actual object is missed by the detection system. This may lead to the increased risk of fatality of the patient. *Selectivity*, on the other hand, refers to the ability to reject artefacts (structures which have some semblance with the object of interest) and is characterised by false positive (FP) rate. False positive detection is undesirable because it leads to unnecessary follow-up procedures which may be expensive.

A desirable detection system, therefore, is one with high TP rate and low FP rate; in other words, good sensitivity and selectivity. In reality, detection systems have limited discriminating abilities and therefore there must be a trade-off between sensitivity and selectivity. The choice between high TP rate and low FP rate depends on prevalence of the disease and the cost of follow-up procedures. If the disease prevalence is rather low, as is the case in microcalcifications, the bias should be towards higher selectivity.

5.3 ROC Analysis

A method called the receiver operating characteristic (ROC) curve analysis is used to validate the detection algorithm developed in this study. The ROC analysis was designed for evaluating the result of signal-detection operations and is used often by radiologists [72]. It is a well-known tool employed in several

microcalcification-detection studies [37, 19, 36, 73, 74, 24, 75]. Hence it can be used to make a rough comparison with other microcalcification-detection techniques.

An explanation of the basic principles of ROC analysis is given by Metz [71] and will be summarised here. Signal-detection processes often involve subjective analysis by the operator (the person performing the detection process) to determine whether the discriminating features of the object of interest are present in the objects being evaluated. Because of the subjectivity involved, the outcome of a detection process has varying degrees of 'confidence levels' or ratings. For example, the confidence levels can be expressed in five degrees: (1) definitely or almost definitely negative, (2) probably negative, (3) possibly positive, (4) probably positive, and (5) definitely or almost definitely positive.

The operator must then establish a 'decision threshold' to separate the true signals from the false ones. In the five-degree confidence level example, if the threshold is set to 4.5, then only signals with confidence level of 5 (definitely or almost definitely positive) will be considered as positive detection, and the rest will be considered as negative detection. Alternatively, if the threshold is set to 1.5, then only the signals with the lowest confidence rating (1) will be considered as negative detection, while the rest become positive detection. Any intermediate level can be chosen as the threshold level, depending on several factors such as the operator's 'style', estimate of the prior probabilities, and so on. The concept of confidence ratings and decision threshold is illustrated in Figure 5.1.

Assume that in a hypothetical signal-detection process, an operator attempts to classify a number of test objects by giving them continuous confidence ratings, based on the operator's chosen method of classification but without knowing the actual status of the objects. If the frequencies of occurrence of signals are plotted with respect to their confidence ratings (*i.e.*, a histogram plot), separating the actually positive objects from the actually negative ones, the result might look like the two curves in Figure 5.1. If a decision threshold is defined as in the figure, then all signals with ratings higher than the threshold would be considered as true (positive detection), whereas those with lower ratings would become negative detection.

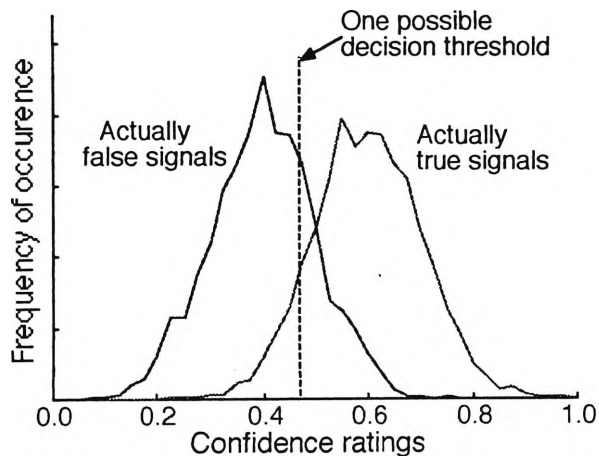

Figure 5.1. Hypothetical frequency distributions of true and false signals with respect to their confidence ratings from zero (definitely false) to one (definitely true), with a possible decision threshold defined. In this situation, all signals to the right of the threshold are detected as positive, while those to the left are considered negative.

If the decision threshold is set very low, all the true signals would be detected (*i.e.*, high TP rate), but the selectivity will be low since many false signals will be also detected (*i.e.*, high FP rate). On the other hand, if the threshold is set very high, both TP and FP rates will be very low. As the threshold is increased, both TP and

FP rates will decrease, thus the sensitivity decreases while selectivity increases. If the confidence rating is based on a properly selected discriminating feature, the FP rate should decrease faster than the TP rate. Plotting the FP-TP pair for every threshold level (or 'operating point') results in the ROC curve. Figure 5.2 shows some examples of ROC curves.

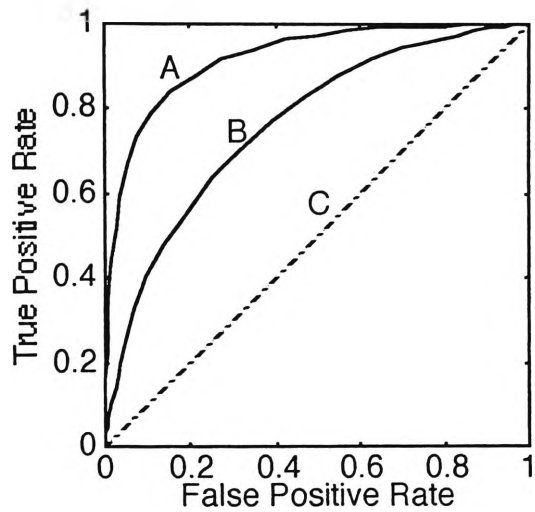

Figure 5.2. Examples of ROC curves.

The ROC analysis can be used for two purposes:

(a) Determining the best threshold level to achieve the optimum balance between sensitivity and selectivity;

(b) Evaluating the effectiveness of the discriminating feature as a detection criterion.

A ROC curve which runs upward to the right along a straight line indicates that the detection is based on a 50-50 chance (Curve C in Figure 5.2). A good ROC curve should have a convex shape towards the upper left corner, which indicates that the detection is based on a good selection criteria, since the TP rate increases

faster than the FP rate (Curves A and B in Figure 5.2). The effectiveness of the selection criteria can be measured from the area under the ROC curve, often denoted by A_z, where larger value generally indicates better performance. In the above example, Curve A is better than Curve B.

Automated detection systems such as those based on neural network or fuzzy logic can readily use the ROC curve, since their outputs are expressed as multi-scale probabilities of the detection process being positive or negative.

ROC analysis is used in this study because the 'gold standard' used in the validation provides only binary information about the status of the lesions in an image. A description about the truth files used as the gold standard is presented in the next section.

5.4 Comparative Analysis

5.4.1 Difference-image Method

As an example of a non-fuzzy microcalcification detection system, the difference-image method as reported by Nishikawa [10] will be used for comparison with the fuzzy detection system. The difference-image method works in the following manner. First, enhancement of small structures and suppression of background texture are performed. This is followed by a global thresholding to remove the remaining background texture. The global threshold level is based on the gray-level histogram of the whole image and is chosen so that 98% of all the

pixels in the difference image are considered as background. After a noise elimination process, a local thresholding is performed. The local threshold level is determined from the mean and standard deviation of the gray-level values of pixels within a 51×51-pixel area of the processed image. The local threshold is defined as a multiple of the local standard deviation above the local mean and can be varied by changing the number of multiples.

The size of the convolution filters for the difference-image method were modified to adjust with the pixel resolution of the images used for validation. In Nishikawa's test, the pixel size of the images employed was 100μm, whereas the images used in this study have a pixel size of 35μm.

5.4.2 Binomial Two-sample Test

The binomial two-sample test method is used to provide a statistical comparison of the performance of the fuzzy detection system and the difference-image method. The two-sample test is appropriate when two different methods are applied to similar subjects [76] and the output of each treatment is binomial. In principle, the two-sample test compares p_x and p_y, the true probabilities of success of methods X and Y, respectively, when method X is applied n times with a success rate of x, and method Y applied m times with a success rate of y. To determine whether one method has a significantly higher success rate than the other, the following expression is calculated [76]:

$$\frac{\frac{x}{n}-\frac{y}{m}}{\sqrt{\frac{\left(\frac{x+y}{n+m}\right)\left(1-\frac{x+y}{n+m}\right)(n+m)}{nm}}} \qquad [5.1]$$

If the result of the above expression is $\leq -z_{\alpha/2}$ or $\geq z_{\alpha/2}$, it can be concluded that one treatment has significantly higher success rate. Otherwise, there is not enough evidence to consider one treatment to have higher success rate than the other. The value of $z_{\alpha/2}$ is given by the Student's t-distribution. At $\alpha = 0.01$, the value of $\pm z_{\alpha/2}$ is ± 2.58.

5.5 Test Data

5.5.1 Mammogram Database

A set of digitised mammograms from a database published by the Lawrence Livermore National Library and the University of California, San Francisco, USA [77] is used for validating the detection system. This database contains 50 sets of mammogram images. Each set represents one patient and contains four images; the images of right and left breasts, each one viewed in two directions (medio-lateral and cranio-caudal). The images are digitised from X-ray films at 35 μm pixel size with 4096 grey levels (12 bits).

The use of this mammogram library has several potential advantages. Since the library is available on CD-ROM and can be acquired publicly, it can be used by

researchers from different institutions studying mammography analysis to compare their methods. Unfortunately, at the time of writing this thesis, there are not that many publications reporting the use of this library in their studies.

The UCSF-LLNL mammogram library has another advantage compared to other publicly available libraries. It has relatively high resolution, both in terms of spatial resolution (*i.e.*, having a fine pixel size) and brightness resolution (the number of gray-scale levels). The higher resolution means that the subtle details of the mammograms are better preserved. On the other hand, it requires a higher amount of memory for processing and storage.

The images contain different degrees of lesions. Some images are normal with no calcifications at all; while others have some calcifications, either singular or clustered. A proportion of images with clustered microcalcifications represents malignant cases, and the rest are benign. The malignancy of the identified microcalcifications is confirmed by either biopsy or a number of follow-up screenings.

Images containing lesions are accompanied by *truth files* which mark the lesions. There are two types of truth files; one which marks the individual calcifications, and the other marks the extent of the cluster area. The truth file which marks the singular lesions, however, only marks a few exemplary microcalcifications and therefore cannot be used as an absolute reference.

5.5.2 Validation Process

In order to evaluate the performance of the detection system, the actual state of the objects in the sample images must be first established as a reference. Ideally, all actual positive lesions must be identified beforehand to avoid a "false-false-positive" (a detection result which is deemed as a false-positive, but is in fact a true positive because the object in question is actually positive in the first place). Unfortunately, the truth files from the library only mark a proportion of the singular lesions. For unmarked objects that may be detected, the actual state is determined using the following procedure:

(a) Objects resembling a microcalcification and located within or close to a cluster area, are accounted as microcalcifications;

(b) Singular objects not resembling a microcalcification and located far from a cluster area, are not considered as microcalcifications;

(c) Singular objects resembling a microcalcification but far from any cluster are considered ambiguous objects. They could be real microcalcifications. But even if they are, they may not be significant since they are not part of a cluster. Therefore in the analysis of singular calcifications, they are considered as true positives, whereas in the analysis of clusters, they are not considered as true positives.

The sample objects are then classified into the following four classes:

1: Non-microcalcifications

2: Ambiguous objects

3: Microcalcifications not marked in the truth files

4: Microcalcifications marked in the truth files

5.5.3 Detection Targets

The target for the detection and the determination of true positive/false positive rates will be defined at three levels:

- Singular calcification level in which the objects are evaluated individually.
- Cluster level where a cluster is considered positive if at least two objects are detected within a distance of 100 pixels, which is equivalent to 3.5mm.
- Image level at which an image is considered normal or negative if no abnormal cluster is detected.

Evaluation at singular calcification level is intended to measure the true performance of the fuzzy operator as a microcalcification detection system, since they work on individual objects without considering the inter-objects characteristics. Meanwhile, the higher levels of evaluation measure the performance of the detection system in practical context. It is possible that some false positive objects detected at singular level do not appear as clusters. This leads to a better performance at cluster level than at singular level. On the other hand, it is also possible that the system has high sensitivity at the individual level but fails to

detect some clusters because only one microcalcification from each of these clusters is detected.

For the singular test, a total of 818 sample objects were tested. These samples were extracted from 41 images and comprise of true microcalcifications (as determined by the truth files and from personal evaluation), ambiguous objects, and false objects which score highly with the peak detector. The TP and FP rates are measured relative to the number of true microcalcifications and non-microcalcifications, respectively, in the set of the test objects.

For the cluster test, 30 images were used, of which only 12 actually contain clusters of microcalcifications. The TP rate for the cluster test is expressed as the percentage of the known clusters which are correctly identified, while the FP rate is expressed as the number of FP clusters in each image. The difference is due to the fact that the number of actually false clusters cannot be determined.

5.6 Results and Comparison

5.6.1 Validation of Fuzzy Operators

A series of experiments was carried on a 500×500-pixel segment to evaluate the performances of the fuzzy peak detector and fuzzy edge detector separately, and also to investigate the factors which influence the performance of the peak detector.

The segment was taken from an image with the code name AKRCC and contains a cluster of clearly visible microcalcifications. The candidate pixels in this segment are determined from the local peaks. There are 491 such pixels, all of which are evaluated by the fuzzy peak detector.

5.6.1.1 Fuzzy Peak Detector

The output of the fuzzy peak detector is shown in Figure 5.3 as histograms of the true and false objects with respect to the membership values assigned by the peak detector.

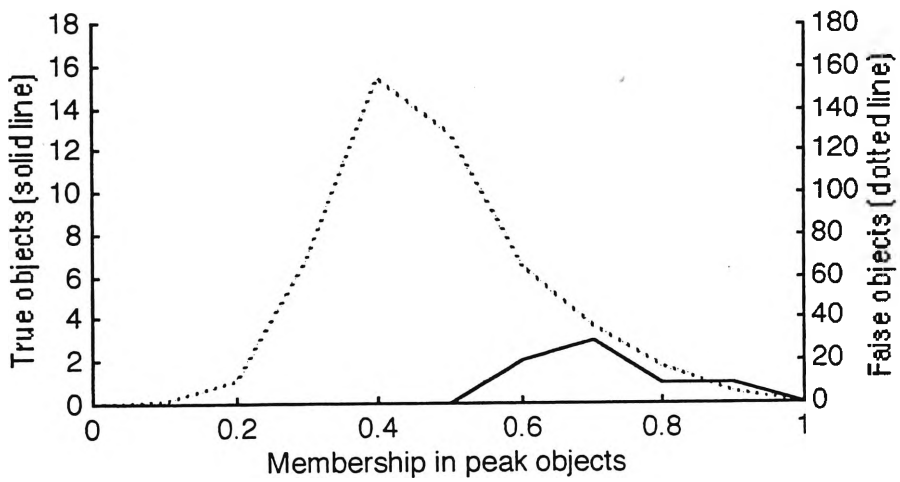

Figure 5.3. Frequency distribution of candidate points with respect to the output from fuzzy peak detector (μ_{peak}) for true objects (left y-axis) and false objects (right y-axis).

It can be seen that the true objects have higher membership values than the majority of the false objects. However, the distributions of the true and false objects are not sufficiently separated; in fact, the distribution of the true object distribution

is almost entirely overlapped by that of the false objects. Therefore, the output of the peak detector alone is not sufficient as a discriminating feature.

To investigate the influence of the fuzzy membership functions on the performance of the peak detector algorithm, different combination of membership functions for *bright*, *dark* and *near* are tested and the results are shown in Appendix A. These results show that, in general, the peak detector's scores for true objects are higher than those for the majority of the false objects, with varying degree of separations between the two classes of objects. Different membership functions produce differences in the absolute values of the peak detector's output. However, in relative terms, the distribution patterns for both classes of objects are the same.

From this test, it can be concluded that changing the membership functions for the fuzzy variables associated with the peak detector only has minor influence on the performance of the peak detector, due to two reasons:

- Slight differences in the membership functions do not affect the outcome of the detection, as far as the separation between the two classes of objects is concerned;

- With the membership functions giving the best separation between the true and false objects, the two classes are still not clearly separated from each other.

5.6.1.2 Fuzzy Edge Detector

From the histogram of μ_{peak} shown in Figure 5.3, it can be seen that the true objects are distributed between 0.5 and 1, with respect to the membership values assigned by the fuzzy peak detector. Because the distribution of the false objects overlap that of the true objects, all pixels within this range of membership values given by the fuzzy peak detector (*i.e.*, higher than 0.5) are processed with the fuzzy edge detector and the result is shown in Figure 5.4.

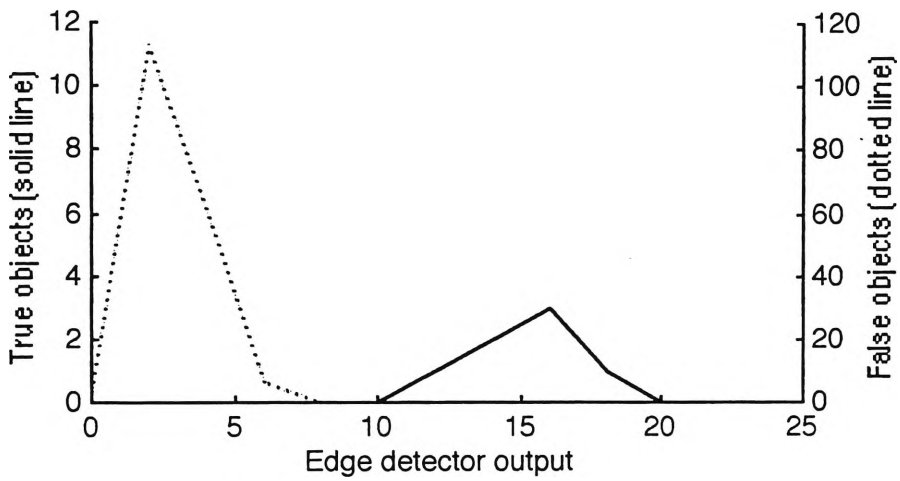

Figure 5.4. Distribution of the output from fuzzy edge analysis (edge strength).

The result of the fuzzy edge detector shows that it improves the result of the fuzzy peak detector. The false objects which were not well separated from the true objects by the peak detector, are now well separated. Considering this outcome, the output of the fuzzy edge detector (*i.e.*, the edge strength) will be used as the discriminating feature in the ROC analysis of the fuzzy detection system.

5.6.2 ROC Analysis of Test Data

The overall results of the tests with the data set as described in Section 5.5.3 are summarised as follows.

5.6.2.1 Singular Microcalcifications

The ROC curve for the singular microcalcifications is shown in Figure 5.5. The area under the curve, A_z, for this curve is 90.65%. At the operating point with the highest gradient change, the TP rate is 77% and the FP rate is 6.6%.

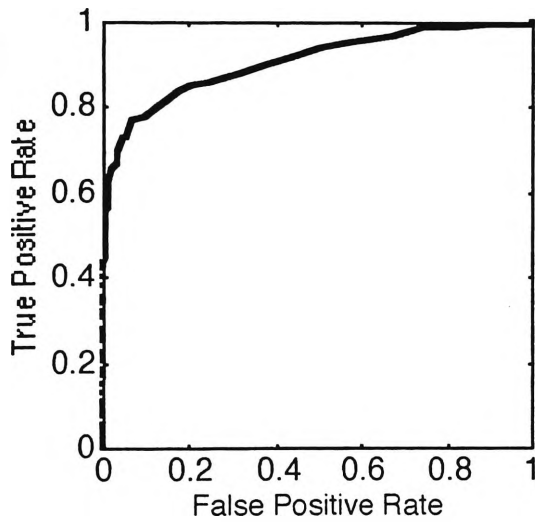

Figure 5.5. ROC curve for singular microcalcifications detection. Both rates are shown as fractions with respect to the number of true or false cases.

5.6.2.2 Clustered Microcalcifications

The ROC curve for the cluster analysis is shown in Figure 5.6. When the FP rate is normalised to the maximum value, the value of A_z for this curve is 90.2%. At

the operating point with the highest gradient change, the TP rate is 87.5% and the FP rate is 0.3 clusters per image.

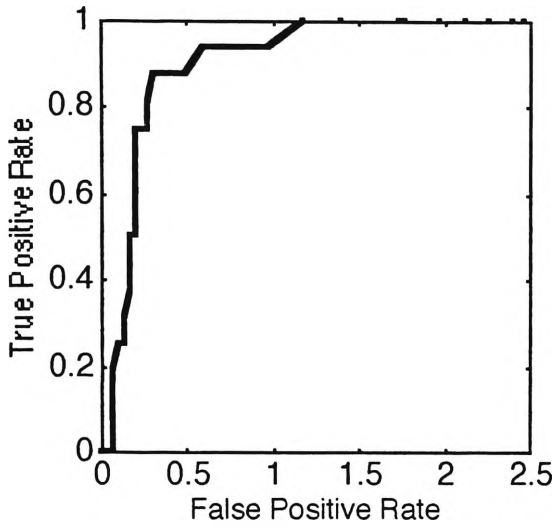

Figure 5.6. ROC curve for clustered microcalcifications detection. The False Positive rate is expressed in number of clusters per image.

5.6.2.3 Full Image

The ROC curve for the image analysis is shown in Figure 5.7. The value of A_z for this curve is 71.99%. At the operating point which gives the highest gradient change in the ROC curve of the cluster analysis, the TP rate is 91.67% and the FP rate is 38.8%.

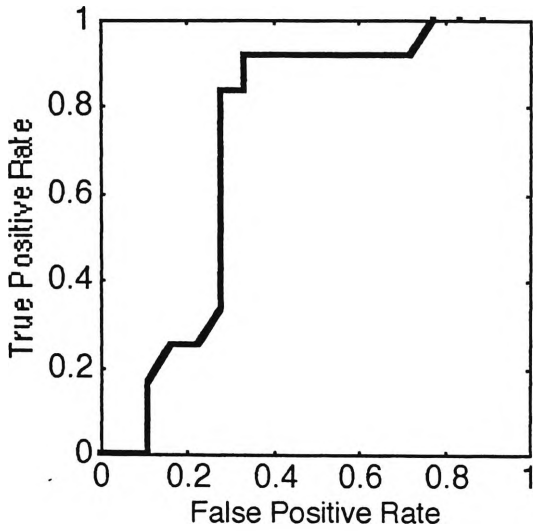

Figure 5.7. ROC curve for whole image evaluation.

5.6.3 Comparison with Difference-Image Method

Nishikawa reported that his detection system has a success rate of approximately 85% true clusters with an average of 2 false-positive clusters per image. These figures alone suggest that the fuzzy method performs better than the difference-image method.

The ROC curve was applied to the output of the difference-image method and the result is shown in Figure 5.8. The area below the curve is 82%.

The area below this ROC curve has less information value than the one for fuzzy system because there is a large gap between the operating point for which TP = 1 and FP = 1, and the next point below that (TP = 87.73%, FP = 25.39%). This is due to the fact that a large proportion of both true objects and false objects are already missed due to the global thresholding. The curve also has a rather

straight shape, which indicates that there is no separation between the true objects and the false objects.

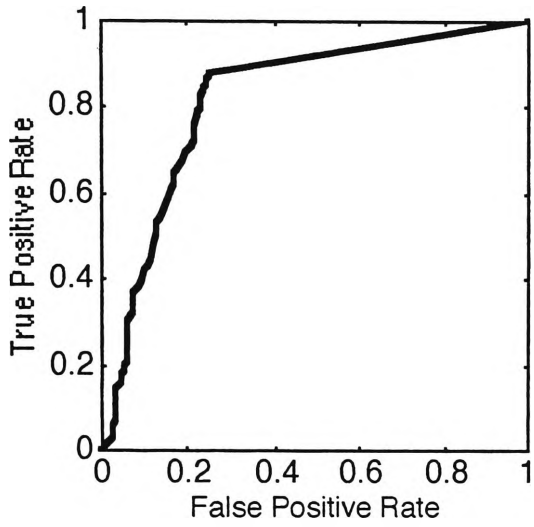

Figure 5.8. ROC curve for the output of difference-image method.

To make a static comparison between the fuzzy and difference-image methods, the accuracy of each is calculated at an operating point which has a similar TP rate for both methods. This is done at two points. The first point is where the difference-image method has the highest TP below 1, which means that the threshold is set just above the lowest possible level. At this operating point, the threshold of the difference-image method is effectively determined by the global threshold, rather than the local threshold. The second point is where the curve of the fuzzy method has an approximately 45° gradient, which is roughly the optimum point for the fuzzy method. The values corresponding to these operating points are tabulated below.

Table 1. Success rates for the Fuzzy and Difference-image methods

	First Operating Point			Second Operating Point		
	TP rate	TN rate	Overall	TP rate	TN rate	Overall
Fuzzy Method	232/266 (0.8722)	335/453 (0.7395)	567/719	224/266 (0.8421)	392/453 (0.8653)	616/719
Difference Image method	236/269 (0.8773)	338/453 (0.7461)	574/722	226/269 (0.8401)	344/453 (0.7594)	570/722

The result of the statistical test as given in Equation 5.1 is tabulated below.

Table 2. Result of statistic test of the two methods

First Point	Second Point	$z_{0.005}$
0.3	3.3	2.58

The statistical test at the first operating point suggests that the two methods have similar accuracy rates. However, the performance of the difference-image method quickly deteriorates as demonstrated by the second operating point, where it has significantly lower accuracy rate.

5.6.4 Comparison with Truth File

The performance of fuzzy method is compared with the objects tagged by the truth files (type 4 objects) and the set of objects tagged definitely normal (type 1 objects). The ROC curve for this test is shown in Figure 5.9. The result is very similar as when objects of types 2 and 3 are included.

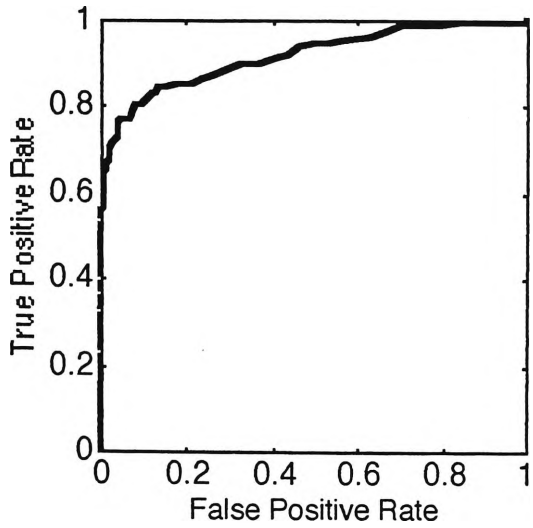

Figure 5.9. ROC curve for test with the objects defined by truth files.

5.6.5 Comparison with Other Edge Detection Methods

To compare the fuzzy edge detection method with conventional edge-detection methods, further tests were conducted by replacing the fuzzy edge detector with three different edge detection methods; Sobel, Prewitt and Roberts. These are implemented in Matlab™ Toolbox. According to the documentation, the Sobel and Prewitt operators are sensitive to horizontal and vertical edges, while the Roberts operator is sensitive to diagonal edges. All methods use the maximum of first derivative to find the edges, similar to the fuzzy method used here. The difference is that the convolution mask for the fuzzy operator is designed to be equally sensitive to diagonal edges as it is to vertical and horizontal edges. The fundamental difference is that the threshold is defined as a crisp threshold in the conventional methods.

ROC curve for the tests with Sobel, Prewitt and Roberts operators are shown in Figures 5.10-5.12. Generally they have lower performance index compared to the

fuzzy system. The areas under the curves are 85% for Sobel, 85% for Prewitt, and 89% for Roberts method compared to a performance of 90.65% for the fuzzy method.

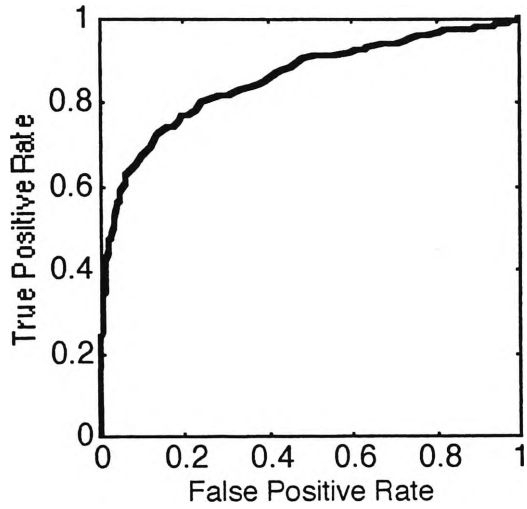

Figure 5.10. ROC curve for detection with Sobel operator. Az = 85%.

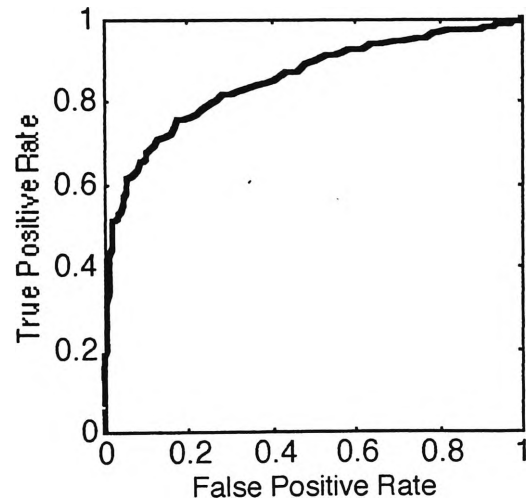

Figure 5.11. ROC curve for detection with Prewitt operator. Az = 85%.

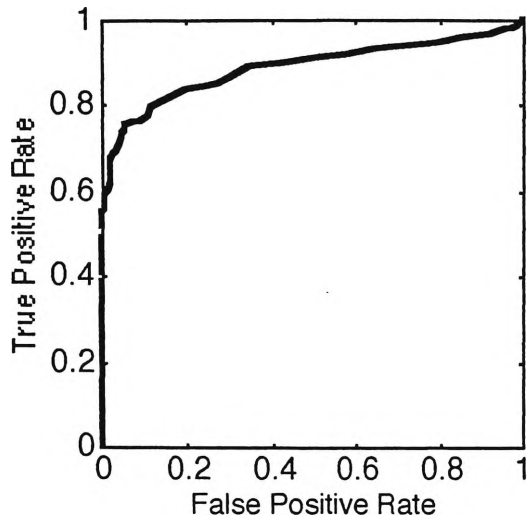

Figure 5.12. ROC curve for detection with Roberts operator. Az = 89%.

5.7 Analysis of Results

The performance of the fuzzy detection system is quite acceptable at the singular level, which is indicated by the value of A_z for that ROC curve. Comparison with the difference-image method also suggests the fuzzy method has a superior performance. At the cluster level, the performance of the two methods is similar. However, the performance of the fuzzy method deteriorates significantly at the image level. The reason for this discrepancy is that at the same operating point, the FP rate for cluster is about 0.3 per image, or approximately 1 FP image in every 3 images, which is roughly the same as the FP rate of the image (38.8%).

5.8 Limitations of the Detection System

A significant proportion of the false negative detection at singular level can be attributed to clusters of very fine calcifications, which appear in a few of the test

images. These calcifications can hardly be recognised as genuine lesions when examined individually, as they appear like common background tissue in the normal images. Their specific characteristics can only be recognised when they are examined collectively. Detection of this type of clustered calcifications requires a more comprehensive analysis of the relationship between the microcalcifications in a cluster.

5.9 Summary

The fuzzy microcalcification detection system has been tested and validated. It can be concluded from the results that the peak detector, together with the edge detection system, are able to detect the microcalcifications in general. However, there is a particular type of microcalcifications which cannot be detected, and would require a different approach. At the image level, further work is required to reduce the false negative rates produced by the technique.

Chapter 6: Conclusion

6.1 Primary Outcomes

Early detection of breast cancer is quite important, as there is no effective method for either prevention or treatment of tumours detected at an advanced stage. In order to reduce mortality rate, mass screening programs are conducted, which require an effective and efficient diagnosis system. To this date, a practical and fully automated mammogram analysis system has not been used on a wide scale and many different methods are still being researched towards reaching that goal.

In the work reported in this thesis, the application of fuzzy image processing in diagnosis of abnormalities in a mammogram has been investigated. Fuzzy logic is one of the emerging methods of soft computing which has been applied in many areas of system automation, control, and also medical decision support systems. It has also been employed in image processing at both low and high levels.

In this work, fuzzy logic has been applied to the analysis of digitised mammograms at the lowest level of image processing to detect microcalcifications. A fuzzy algorithm has been developed to detect a microcalcification in a

mammogram by a fuzzy peak detector, to identify its edge using a fuzzy edge detector, and to define the edge by a fuzzy function.

6.2 Effectiveness of the Algorithms

The fuzzy peak detector has a simple algorithm, since it only uses membership functions to analyse the pixels in the neighbourhood of a candidate pixel. Although there is no complex mathematical function utilised, the process is quite computing intensive. This is due to the structure of digital computers which is based on binary logic.

The fuzzy edge detector, on the other hand, is slightly more complex than the conventional edge operators as it employs a set of non-linear fuzzy rules and membership functions. It has been shown in this work that the fuzzy edge detector has a superior performance compared to the conventional methods.

The sensitivity of the fuzzy detection method for individual microcalcifications has proved superior compared to the other methods reported in the literature. On the other hand, the sensitivity of the fuzzy method for clusters of microcalcifications has not been as good as of singular microcalcifications. This is due to presence of small microcalcifications with low contrast and diffused boundaries in a cluster. Such microcalcifications cannot be detected even individually.

The application of the fuzzy algorithms to a complete image to classify it as normal or suspicious has produced a poor performance clinically. A rate of 38.8%

false positive (which is, roughly, 2 false positive images in every 5 images) is too high for the system to be used clinically. Clearly the selectivity of the system needs improvement.

Comparison with a non-fuzzy detection method; the difference-image method, the fuzzy system has shown a better performance. The poor performance of the difference-image method has been due to the application of global thresholding which fails to detect weak signals.

It should be also pointed out that the test data employed in this study has a higher prevalence rate of lesions than usually encountered in actual clinical screening program. This has adversely affected the performance of the algorithm. Hence in real clinical work, the rate of false positive detection reported will decrease when the algorithms are applied to mammograms obtained in screening programs.

6.3 Limitation of Work

The fuzzy peak detector fails to detect some clusters marked in the truth files. In visual observation, those missed clusters appear to contain very fine but dense grains of microcalcifications. The individual grains have little resemblance to a microcalcification and visually they can be identified as calcifications when there are similar recognisable microcalcifications in their vicinity. While the fuzzy peak detector is relatively insensitive to slight differences in size and contrast of peak

objects, it is not able to detect such fine microcalcifications. Such a failure is due to two reasons:

(a) The fuzzy detector is modelled based on the characteristics of singular microcalcifications relatively larger and brighter than the fine microcalcifications.

(b) The fuzzy peak detector operates only on the pixels located within the window of a local peak. Since the distance between calcifications in a cluster is larger than the window size, the presence of other objects cannot be determined or assessed by the peak detector.

Another limitation of the developed algorithms is its inability to determine whether a detected microcalcifications is malignant or benign. This is because the algorithms are currently only pixel-based and do not take into account more regional or global features of an image such as geometrical and textural information which depend on the relationship between two or more microcalcification.

6.4 Recommendation for Future Research

The research work reported in this thesis has contributed to increase understanding in the application of fuzzy theories in medical image processing, and has produced encouraging results, even in limited testing conditions. However, a considerable additional work is required before it can be proven as an effective solution to the problem stated in the introduction of this thesis. The aspects of the proposed system that need improvement are execution speed of the algorithm,

sensitivity to fine microcalcifications, ability to assess the malignancy of detected lesions, and involvement of local experts in the validation process.

The fuzzy algorithms have been developed in this study using Matlab™ functions. As an interpreted environment, Matlab™ has limited processing power in terms of execution speed and the size of data it can handle. In order to increase the execution speed of the algorithms, it is necessary to optimise the code by compiling the functions using a suitable compiler.

The performance of the algorithms for fine clustered microcalcifications can be improved by reducing the size of the window used for the peak analysis. Alternatively, a secondary peak operator can be introduced, designed based on small lesions. The result can be also combined with the textural information of the local and regional areas to produce a better detection rate.

The detection system can be expanded to analyse the malignancy of the detected clusters. Apart from the density of the clusters (*i.e.*, the number of grains in a certain area), another indication of malignancy is the shape of the grains, their configuration and orientation. This can be achieved by expanding the edge detection algorithm to carry out the geometrical analysis of the edges.

Although the mammogram library employed in the validation is of excellent quality, it can be further improved by supplementing it with local experts' opinion for ambiguous microcalcifications.

Bibliography

[1] L. O. Hall, "Learned fuzzy rules versus decision trees in classifying microcalcifications in mammograms," *Proc. SPIE Applications of Fuzzy Logic Technology III*, Bruno Bosacchi; James C. Bezdek; Eds. Vol. 2761, p. 54-61, 1996.

[2] S. Bothorel, B. B. Meunier, S. A. Muller, "Fuzzy logic based approach for semiological analysis of microcalcifications in mammographic images," *International Journal of Intelligent Systems*. 12(11-12):819-848, Nov-Dec 1997.

[3] http://hna.ffh.vic.gov.au/phb/hdev/canhrtag/breast.html

[4] J. E. Martin, M. Moskowitz, J. R. Milbrath, "Breast cancer missed by mammography", *AJR*, pp. 132-737, 1979.

[5] L. Kalisher, "Factors influencing false negative rated in mammography", *Radiology*, pp. 133-297, 1979.

[6] B. J. Herman,et al, "Occult Malignant breast lesions in 114 patients: relationship to age and the presence of microcalcifications", *Radiology*, pp. 169-321, 1988.

[7] S-C. B. Lo, H. Li, J-S. Lin, A. Hasegawa, O. Tsujii, M. T. Freedman, S. K. Mun, "Detection of clustered microcalcifications using fuzzy modeling and convolution neural network", *Proc. SPIE Medical Imaging 1996: Image Processing*, Murray H. Loew; Kenneth M. Hanson; Eds. Vol. 2710, p. 8-15, 1996.

[8] S. Astley, I. Hutt, S. Adamson, P. Miller, P. Rose, C. Boggis, C. Taylor, T. Valentine, J. Davies, and J. Armstrong, "Automation in mammography: computer vision and human perception," *State of the art in digital mammographic image analysis*, ed. K. W. Bowyer and S. Astley, World Scientific, Singapore, 1994.

[9] J. Dengler, S. Behrens, and J. F. Desaga, "Segmentation of microcalcifications in mammograms," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp 634-642, 1993.

[10] R. M. Nishikawa, M. L. Giger, K. Doi, C. J. Vyborny, R. A. Schmidt, "Computer-aided detection of clustered microcalcifications on digital mammograms," *Medical & Biological Engineering & Computing*, no 33, pp 174-178, 1995.

[11] D. Nesbitt, F. Aghdasi, R. Ward, J. Morgan-Parkes, "Detection of microcalcifications in digitized mammogram film images using wavelet enhancement and local adaptive false positive suppression," *IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing. Proceedings*, 1995.

[12] I. N. Bankman, W. A. Christens-Barry, I. N. Weinberg, D. W. Kim, R. D. Semmel, W. R. Brody, "An algorithm for early breast cancer detection in mammograms," *Proceedings. Fifth Annual IEEE Symposium on Computer-Based Medical Systems*, 1992.

[13] L. A. Zadeh, "Knowledge representation in fuzzy logic". *An Introduction to fuzzy logic applications in intelligent systems*, ed. R.R. Yager, L.A. Zadeh, Kluwer Academic, Boston, pp. 1-21, 1992.

[14] L. A. Zadeh, "Fuzzy Sets", *Information and Control*, v. 8 pp. 338-353, 1965.

[15] A. Kricker, P. Jelfs, "Breast cancer in Australian women 1921-1994: Summary", [online], January 1999,

http://www.nbcc.org.au/pages/info/resource/nbccpubs/bc21-94/summary.htm.

[16] http://www.nswcc.org.au/pages/ccic/ccr/sumbreas.htm

[17] http://www.ozemail.com.au/~glensan/cancer.htm

[18] L. W. Bassett, "Mammographic Features of Malignancy", *The Female Breast and its Disorders*, G. W. Mitchell Jr. and L. W. Bassett (eds.), Williams & Wilkins, 1990.

[19] A. P. Dhawan, T. Chitre, C. Kaiser-Bonasso, M. Moskowitz, "Analysis of mammographic microcalcifications using gray-level image structure features", *IEEE Trans. Med. Imaging*, 15 (3), p. 246-259, 1996.

[20] Sylvia H. Heywang-Köbrunner, "Diagnostic breast imaging: mammography, sonography, magnetic resonance imaging, and interventional procedures", Stuttgart, Thieme, 1997.

[21] R. H. Gold, "Mammographic features of benign disease", *The Female Breast and its Disorders*, G. W. Mitchell Jr. and L. W. Bassett (eds.), Williams & Wilkins, 1990.

[22] G. G. Lee, C. H. Chen, "A multiresolution wavelet analysis and Gaussian Markov random field algorithm for breast cancer screening of digital mammography," *1996 IEEE Nuclear Science Symposium Conference Record*, 1996.

[23] H. D. Li, M. Kallergi, L. P. Clarke, V. K. Jain, R. A. Clark, "Markov random field for tumor detection in digital mammography," *IEEE Transactions on Medical Imaging*, vol. 14, no. 3, p. 565-76, 1995.

[24] H. Yoshida, Wei Zhang, W. Cai, K. Doi, R. M. Nishikawa, M. L. Giger, "Optimizing wavelet transform based on supervised learning for detection of microcalcifications in digital mammograms," *Proceedings. International Conference on Image Processing*, 1995.

[25] J. C. Russ, "The image processing handbook", 2nd ed., CRC Press, 1995.

[26] R. C. Gonzalez, "Digital image processing", Addison-Wesley, 1992.

[27] D. Brzakovic, X. M. Luo, P. Brzakovic, "An approach to automated detection of tumors in mammograms", *IEEE Trans. on Medical Imaging*, v. 9, no. 3, pp. 233-241, 1990.

[28] M. Sameti, R. K. Ward, "A fuzzy segmentation algorithm for mammogram partitioning", *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, R. A. Schmidt, eds; Elsevier Science B.V., pp 471-474, 1996.

[29] D. Brzakovic, N. Vujovic, M. Neskovic, "Early detection of cancerous changes by mammogram comparison," *Proc. SPIE Visual Communications and Image Processing '94*, Aggelos K. Katsaggelos; Ed. Vol. 2308, p. 1520-1531, 1994.

[30] E. A. Stamatakis, I. W. Ricketts, A. Y. Cairns, C. Walker, P. E. Preece, "Detecting abnormalities on mammograms by bilateral comparison," *IEE Colloquium on Digital Mammography*, 1996.

[31] L. N. Mascio, J. M. Hernandez, C. M. Logan, "Automated analysis for microcalcifications in high resolution digital mammograms", [Online], 7 January 1998,

http://www-dsed.llnl.gov/documents/imaging/jmhspie93.html.

[32] H. D. Cheng, M. L. Yui, R. I. Freimanis, "A new approach to microcalcification detection in digital mammograms," *1996 IEEE Nuclear Science Symposium*, 1996.

[33] G. Naghdy, Yue Li, J. Wang, "Wavelet Based Adaptive Resonance Theory (ART) Neural Network for the Identification of Abnormalities in Mammograms," *National Health Informatics Conference*, HIC 97, Sydney, Australia. pp67. Paper 67 on CDROM publication.

[34] H. Li, K. J. R. Liu, S. C. B. Lo, "Fractal modeling of mammogram and enhancement of microcalcifications", *Nuclear Sci. Symp. 1996 Conf. Record, IEEE*, pp. 1850-1854, 1997.

[35] T. O. Gulsrud, S-O. Gabrielsen, "Classification of microcalcifications using a multichannel filtering approach", *Eng. in Med. & Biol. 17th Ann. Conf. 1995 IEEE*, v. 2, pp. 889-890, 1995.

[36] A. Y. Cairns, I. W. Ricketts, D. Folkes, M. Nimmo, P. E. Preece, A. Thompson, C. Walker, "The automated detection of clusters of microcalcifications," *IEE Colloquium on 'Applications of Image Processing in Mass Health Screening'* (Digest No. 056), 1992.

[37] Yulei Jiang, R. M. Nishikawa, D. E. Wolverton, M. L. Giger, K. Doi; R. A. Schmidt, C. J. Vyborny, "Mammographic feature analysis of clustered microcalcifications for classification of breast cancer and benign breast diseases," *Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Engineering Advances: New Opportunities for Biomedical Engineers*, 1994.

[38] F. Aghdasi, R. K. Ward, B. Palcic, "Classification of mammographic microcalcification clusters," *1993 Canadian Conference on Electrical and Computer Engineering*, pp. 1196-1199, 1993.

[39] L. W. Estevez, N. D. Kehtarnavaz, "Computer assisted enhancement of mammograms for detection of microcalcifications", *Proc. 8th IEEE Symp. Computer-based Med. Systems*, pp. 16-23, 1995.

[40] I. N. Bankman, J. Tsai, D. W. Kim, O. B. Gatewood, W. R. Brody, "Detection of microcalcification clusters using neural networks," *Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Engineering Advances: New Opportunities for Biomedical Engineers*, 1994.

[41] N. A. Murshed, F. Bortolozzi, R. Sabourin, "Classification of cancerous cells based on the one-class problem approach," *Proc. SPIE Applications and Science of Artificial Neural Networks II*, Steven K. Rogers; Dennis W. Ruck; Eds. Vol. 2760, p. 487-494, 1996.

[42] N. A. Murshed, F. Bortolozzi, T. Sabourin, "A fuzzy ARTMAP-based classification system for detecting cancerous cells, based on the one-class problem approach," *Proceedings of the 13th International Conference on Pattern Recognition*, 1996.

[43] N.A. Bridgett, J. Brandt, C.J. Harris, "A neurofuzzy route to breast cancer diagnosis and treatment." *Proceedings of 1995 IEEE International Conference on Fuzzy Systems. The International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium*. 1995.

[44] H. Sidaoui, A. de Alburqueque, R. Benjamins, L. Gualberto, "Fuzzy approach with a multi-expert approximate knowledge representation for medical diagnostic", *Systems, Man, and Cybernetics*, v.3 pp. 2290-2293, 1994.

[45] D. L. Hudson, M. E. Cohen, "Fuzzy logic in medical expert systems", *IEEE engineering in medicine and biology*, v. 13/5 pp. 693-698, Nov/Dec 1994.

[46] K. S. Leung, W. Lam, "Fuzzy concepts in expert systems", *Computer*, v. 21/9, pp. 43-56, September 1998.

[47] E. Cox, "The fuzzy systems handbook", AP Professional, 1994.

[48] D. McNeill, "Fuzzy Logic", Simon & Schuster, New York, 1993.

[49] H. R. Tizhoosh, "Fuzzy Image Processing: Potentials and State of the Art", *Methodologies for the Conception, Design and Application of Soft Computing; Proc. IIZUKA '98*, pp. 321-324, 1998.

[50] Z. Chi, H. Yan, T. Pham, "Fuzzy algorithms: with applications to image processing and pattern recognition", World Scientific Publishing Co., 1996.

[51] L. Khodja, L. Foulloy, E. Benoit, "Fuzzy Clustering for Color Recognition Application to Image Understanding", *Proc. 5th IEEE Int. Conf. on Fuzzy Systems*, v. 2 pp. 1407-1413, 1996.

[52] A. L. Ralescu, J. G. Shanahan, "Fuzzy perceptual grouping in image understanding", *Proc. IEEE Int. Conf. on Fuzzy Systems 1995*, v.3 pp. 1267-1272, 1995.

[53] T. Sarkodie-Gyan, C-W. Lam, D. Hong, A. W. Campbell, "A Fuzzy Clustering Method for Efficient 2-D Object Recognition", *Proc. 5th IEEE Int. Conf. on Fuzzy Systems*, v. 2 pp. 1400-1406, 1996.

[54] G. M. Khamseh, "Fuzzy logic based decision support system in pathology", Master's thesis, University of Wollongong, 1997.

[55] S. Zahan, C. Michael, S. Nikolakeas, "A fuzzy hierarchical approach to medical diagnosis", *Fuzzy Systems, 1997. Proc. 6th IEEE Int. Conf. on*, v. 1, pp. 319-324, 1997.

[56] B-T. Chen, Y-S. Chen, W-H. Hsu, "Image Processing and Understanding based on the Fuzzy Inference Approach", *Proc. 3rd IEEE Conf. on Fuzzy Systems*, pp. 254-259, 1994.

[57] C. Demko, E. Zahzah, "Image understanding using fuzzy isomorphism of fuzzy structures," *IEEE Int. Conf. on Fuzzy Systems*. v 3 p 1665-1672, 1995.

[58] C. Bezdek, R. Chandrasekhar, Y. Attikiouzel, "New fuzzy model for edge detection," *Proc. SPIE Applications of Fuzzy Logic Technology III*, Bruno Bosacchi; James C. Bezdek; Eds. Vol. 2761, p. 11-28, 1996.

[59] C-Y. Tyan, P. P. Wang, "Image processing-enhancement, filtering and edge detection using the fuzzy logic approach," *Second IEEE International Conference on Fuzzy Systems*, 1993.

[60] W. Li, G. Lu, Y. Wang, "Recognizing white line markings for vision-guided vehicle navigation by fuzzy reasoning", *Pattern Recognition Letters*, v. 18 pp. 771-780, 1997.

[61] K. H. L. Ho, N. Ohnishi, "FEDGE – Fuzzy Edge Detection by Fuzzy Categorization and Classification of Edges", [online], January 1999,

http://www.bmc.riken.go.jp/sensor/Ho/fedge/fedge.html.

[62] N. W. Granville, "Fuzzy analysis of fuzzy images", *Intelligent methods in health care and medical applications (Digest 1998/514), IEE Coll. on*, pp. 3/1-4, 1998.

[63] J. M. Keller, "Fuzzy Logic Rules in Low and Mid Level Computer Vision Tasks", *Proceedings, 1996 Biennial Conference of the North American Fuzzy Information Processing Society*, pp. 19-22, 1996.

[64] X. Q. Li, Z. W. Zhao, H. D. Cheng, C. M. Huang, R. W. Harris, "A fuzzy logic approach to image segmentation", *Pattern Recognition 1994*, pp. 337-340, 1994.

[65] Y. Zhu, Z. Chi, H. Yan, "Brain image segmentation using fuzzy classifiers", *Intelligent Information Systems, 1994. Proc. 1994 2nd ANZ conf. on*, p 244-247, 1994.

[66] C-W. Chang, G. R. Hillman, H. Ying, T. A. Kent, J. Yen, "A two-stage human brain MRI segmentation scheme using fuzzy logic", *Fuzzy systems, 1995. Proc. 1995 IEEE Int. Conf.*, v.2 pp. 649-654, 1995.

[67] C. Tresp, M. Jäger, M. Moser, J. Hiltner, M. Fathi, "A new method for image segmentation based on fuzzy knowledge", *Intelligence and Systems, 1996. IEEE Int. Joint Symposia on*, pp. 227-233, 1996.

[68] J. M Molina, M. J. Martin, P. Isasi, A. Sanchis, "A fuzzy reasoning system for boundary detection in radiological images", *Fuzzy Systems, 1998 IEEE World Conf. on Computational Intelligence*, v. 2, pp. 1524-1529, 1998.

[69] J. M. Keller, P. Gader, O. Sjahputra, C. W. Caldwell, H-M. T. Huang, "A fuzzy logic rule-based system for chromosome recognition", *Computer-based Medical Systems, 1995. Proc. 8th IEEE Symposium on*, pp. 125-132, 1995.

[70] B. J. Smith, P. Arabshahi, "A fuzzy decision system for ultrasonic prenatal examination enhancement", *Fuzzy Systems, 1996. Proc. 5th IEEE Int. Conf. on*, v. 3, pp. 1712-1717, 1996.

[71] C. E. Metz, "Basic principles of ROC analysis", *Seminars in Nuclear Medicine*, v. 8, No. 4 (October), pp. 283-298, 1978.

[72] S. J. Dwyer III, S. Berr, M. B. Williams, "Modeling mammography examinations", *Computer-based Medical Systems, 1998. Proc., 11th IEEE Symposium on*, pp. 118-131, 1998.

[73] H-D. Cheng, Y. M. Lui, R. I. Freimanis, "A novel approach to microcalcification detection using fuzzy logic technique", *IEEE Trans. on Medical Imaging*, v. 17, no. 3, June 1998, pp. 442-450, 1998.

[74] A. Hojjatoleslami, L. Sardo, J. Kittler, "An RBF based classifier for the detection of microcalcifications in mammograms with outlier rejection capability", *Neural Networks,1997. International Conference on*, v. 3, pp. 1379-1384, 1997.

[75] J. K. Kim, H. W. Park, "Surrounding region dependence method for detection of clustered microcalcifications on mammograms", *Image Processing, 1997. Proc., Int. Conf. on*, v.3 pp.535-538, 1997.

[76] R. J. Larsen, M. L. Marx, An introduction to mathematical statistics and its applications, Prentice-Hall, New Jersey, 1986.

[77] E. A. Ashby, J. M. Hernandez, C. M. Logan, L. N. Mascio, "UCSF/LLNL high resolution digital mammogram laboratory", *Proc. 17th IEEE Int. Conf. Eng. in Med. & Biol. Soc*, pp. 539-540, 1995.

Appendix A: Results of Experiments

A.1 Sensitivity of the Fuzzy Peak Detector to Membership Functions

The following sets of figures (Figure A.1 to Figure A.24) illustrates the results of fuzzy peak detector operations on a set of candidate pixels from a segment of digitised mammogram. Each set of figures displays the membership functions for the fuzzy variables *bright*, *dark*, *near* and *far* on the left hand column. The graph on the right hand column shows the frequency distributions of the true objects and false objects, with respect to the values assigned to them by the fuzzy peak detector. The distributions are plotted with a double y-axis, showing the true objects with solid line on the left hand y-axis, and the false objects with dotted line on the right hand y-axis.

The results that are presented here represent different combinations of membership functions. The membership functions are changed so that they have different supports (the range of input for which the output is non-zero). The supports of a membership function and its opposite determine the magnitude of the object to which the fuzzy detector is sensitive.

To make the detector sensitive to small objects, the membership function for *near* can be given narrow support; to make it sensitive to larger objects, the support should be made wider (cf. Figure A.9 vs. Figure A.8).

With respect to the detector's sensitivity to object's contrast, different membership functions for *bright* and *dark* as the following are tested:

- overlapping vs. non-overlapping supports of *bright* and *dark* (cf. Figure A.2 vs. Figure A.3 vs. Figure A.4)
- wider support vs. narrower support for *bright* (cf. Figure A.8 vs. Figure A.12).

The results shows that although the different membership functions give different degrees of separation between true objects and false objects, in general the true objects have higher membership values than the majority of the false objects.

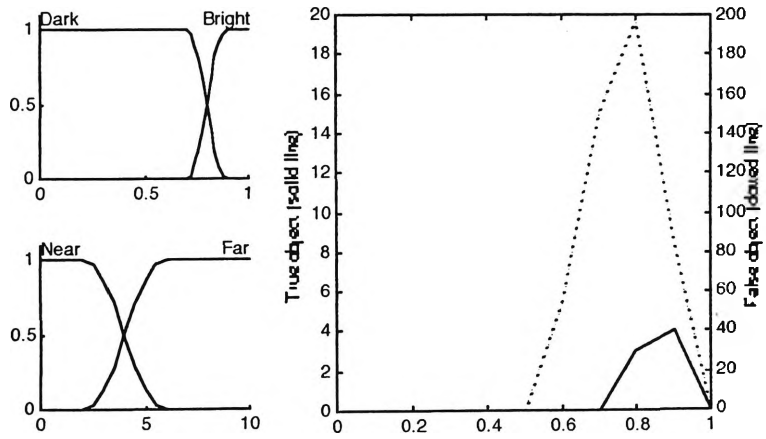

Figure A.1

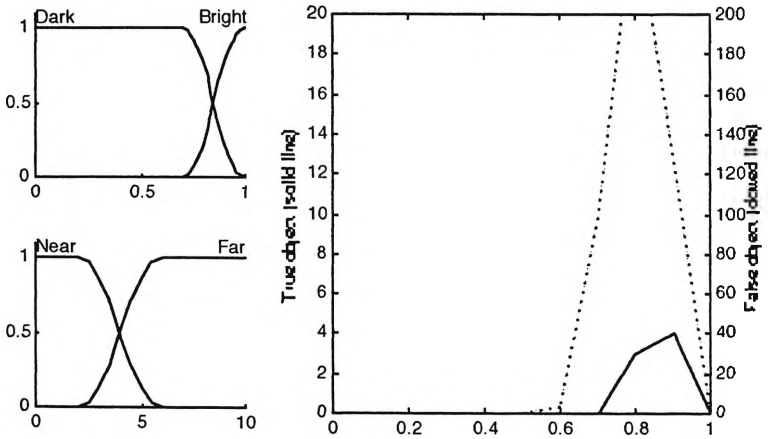

Figure A.2

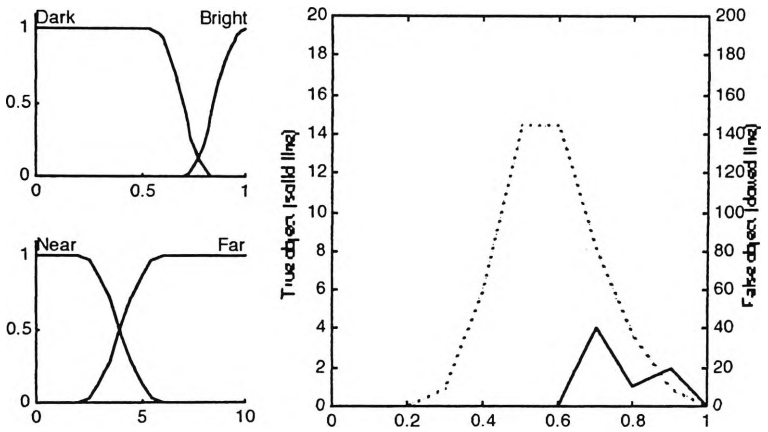

Figure A.3

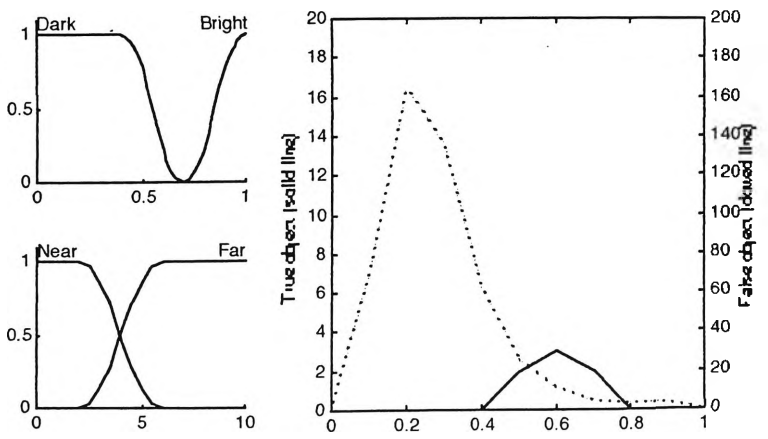

Figure A.4

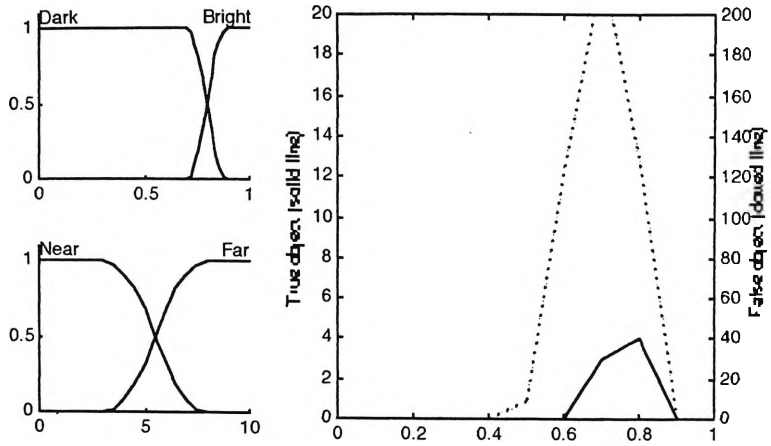

Figure A.5

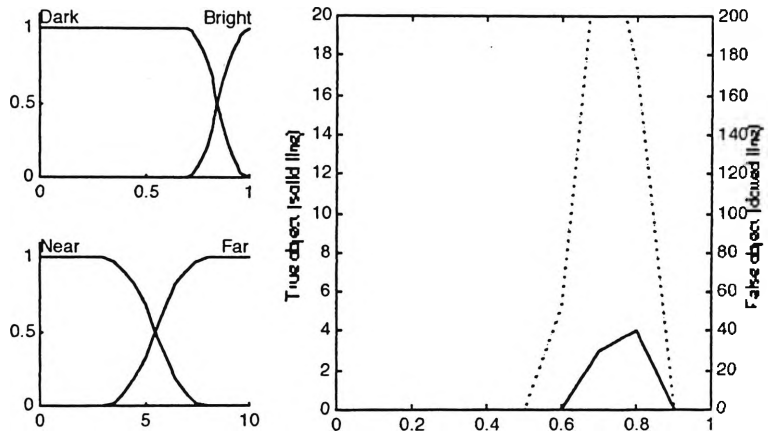

Figure A.6

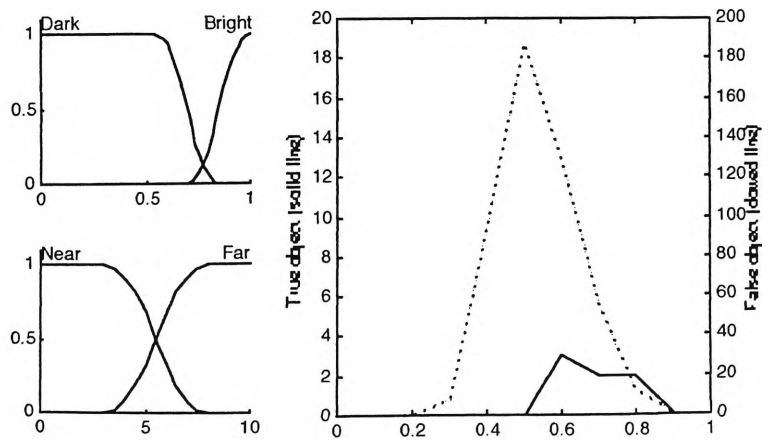

Figure A.7

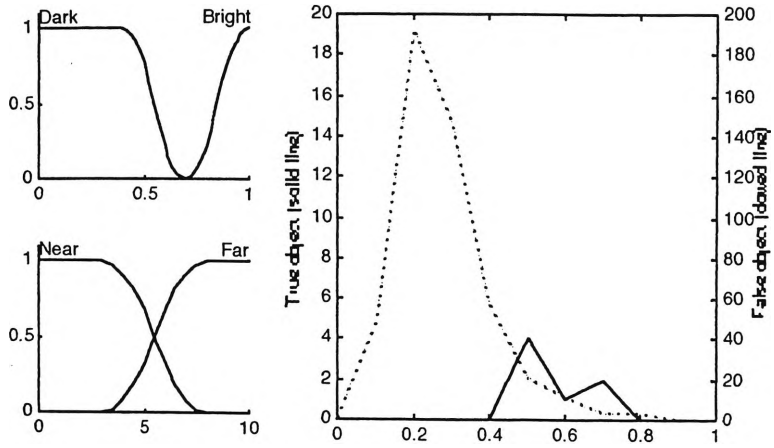

Figure A.8

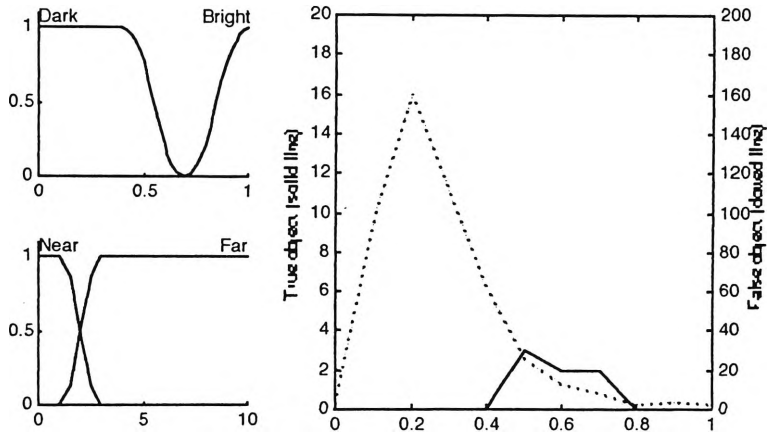

Figure A.9

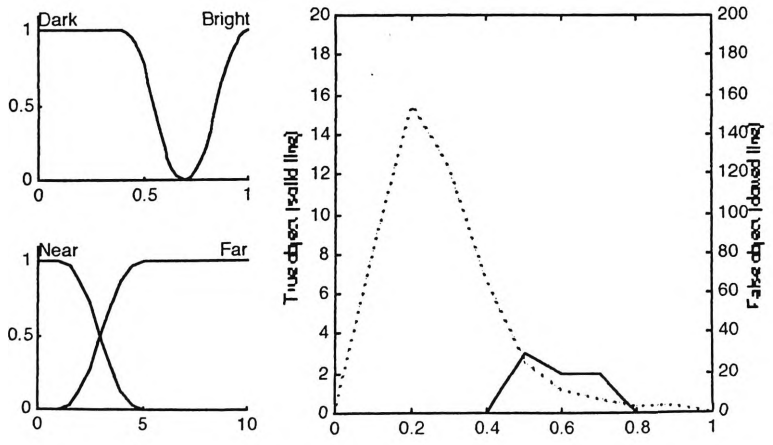

Figure A.10

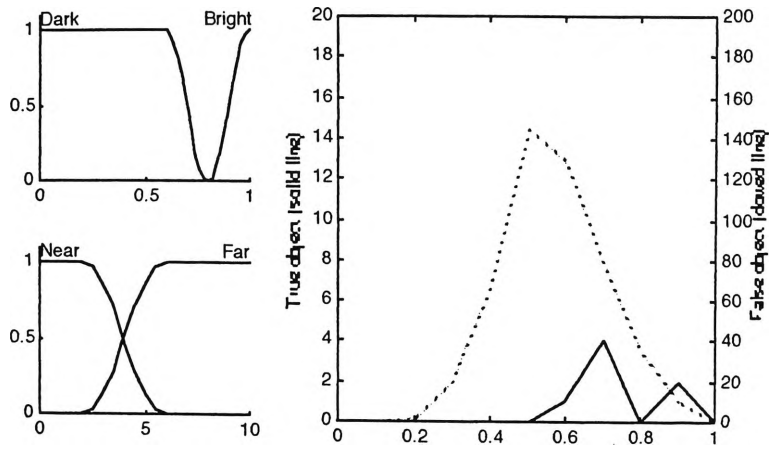

Figure A.11

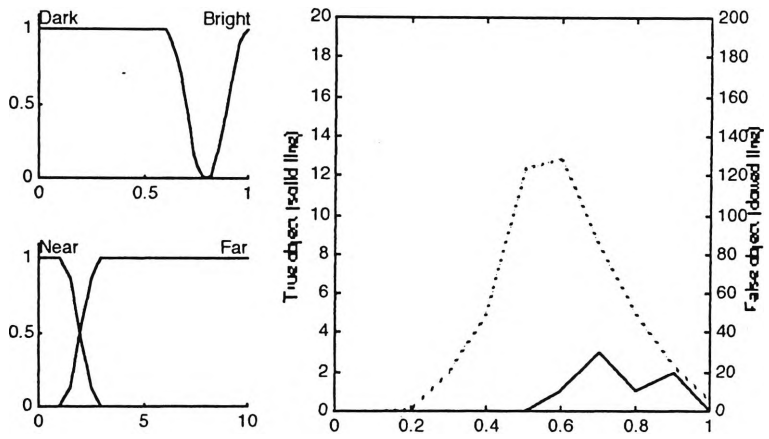

Figure A.12

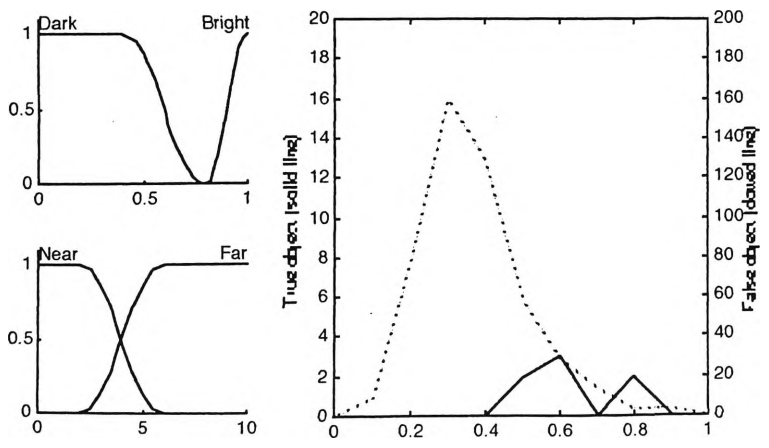

Figure A.13

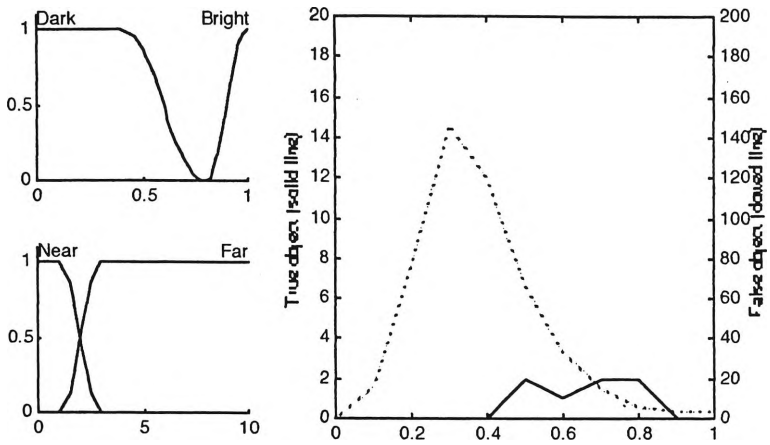

Figure A.14

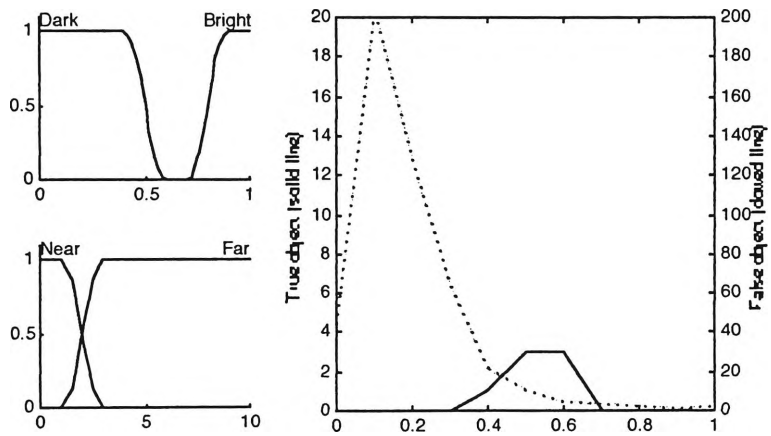

Figure A.15

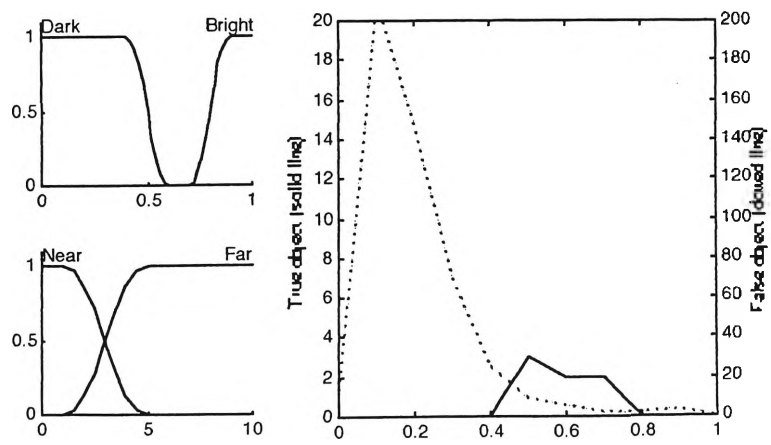

Figure A.16

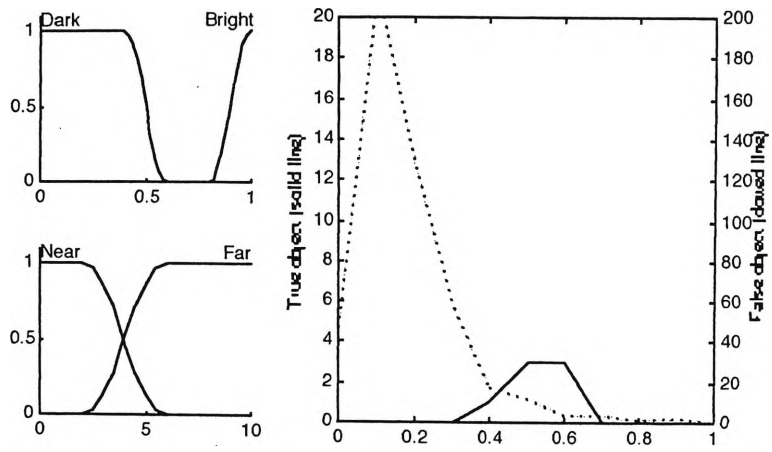

Figure A.17

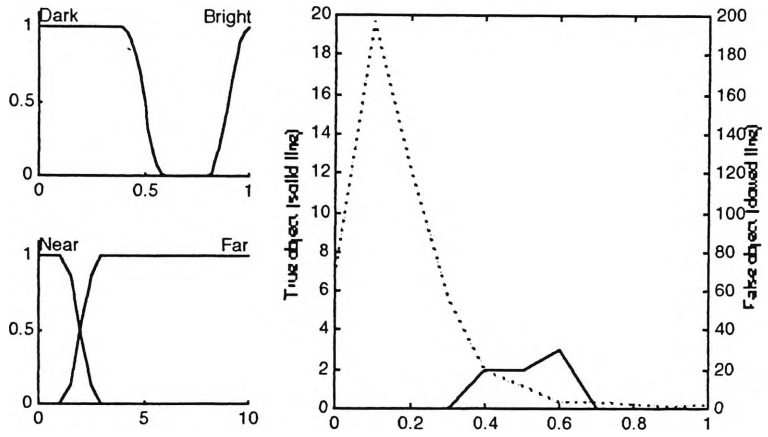

Figure A.18

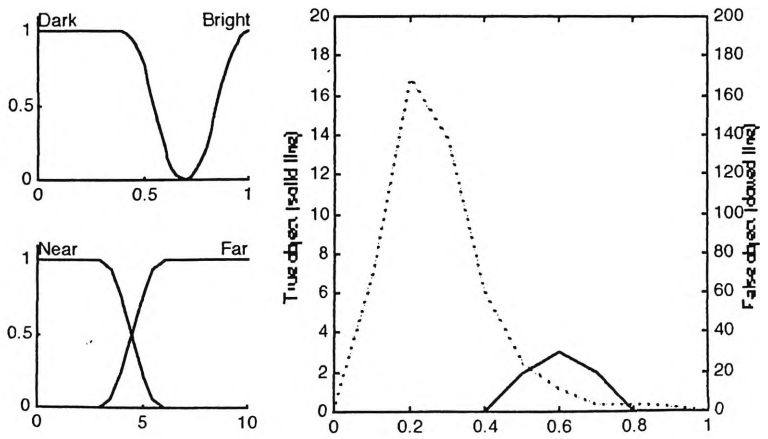

Figure A.19

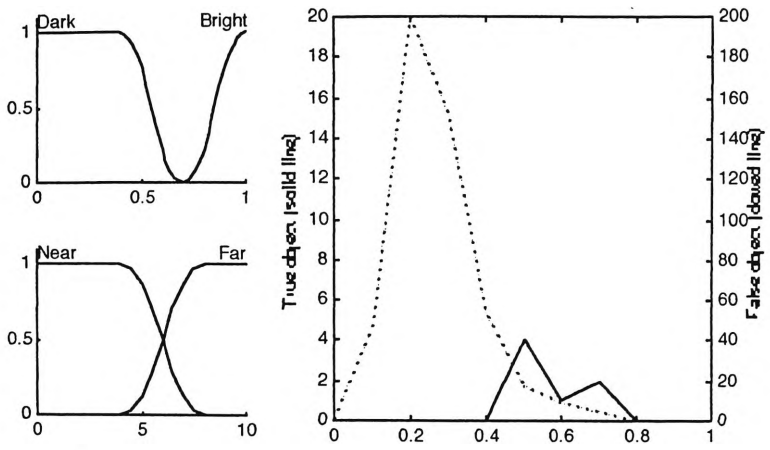

Figure A.20

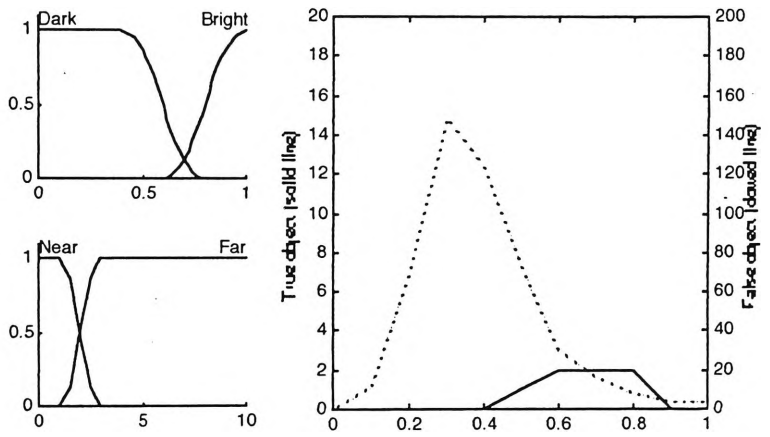

Figure A.21

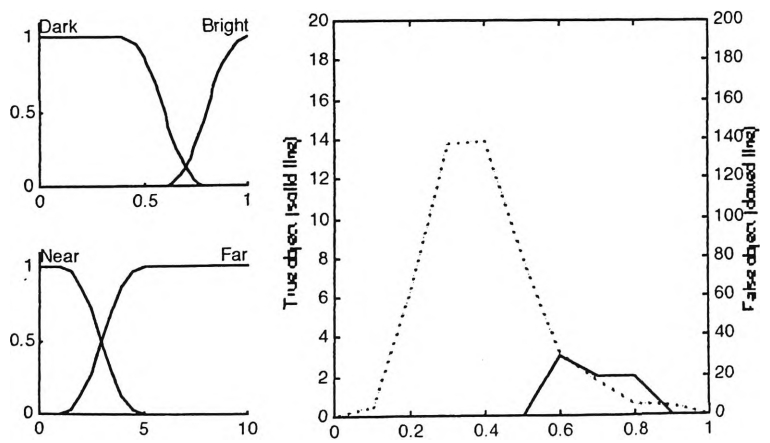

Figure A.22

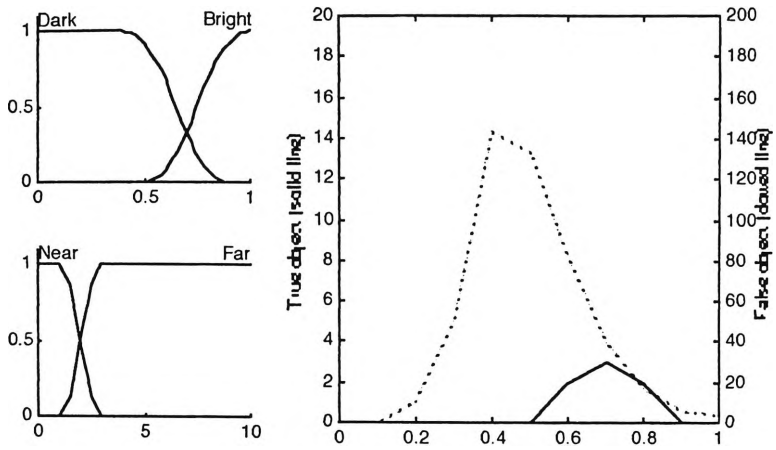

Figure A.23

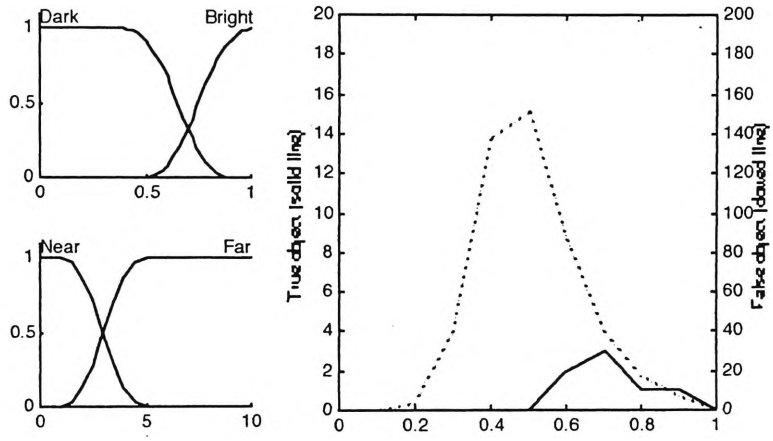

Figure A.24

Appendix B: Program Listing

B.1 Overview

The programs used in this work have been developed and tested in Matlab™ version 5.2 for Windows™, on a Pentium-166 computer with 32 MB of RAM.

B.1.1 Preparing the images

OPENIDS3 prepares the main image by opening the raw .ids file and saving it in Matlab binary format as 500×500 chunks. OPENTRU2 reads the truth files (both the _mc.ids and _cr.ids files) and saves the coordinates of 'true' bits in a vector as Matlab binary format. XYPLOT is used to view the data from the truth files.

B.1.2 Operation on the images

RUN01 prepares the 500×500 segments of the image by adding a portion of the adjacent segment, so that windowing operations can be performed on the segments properly. RUN03 is the batch file to perform the fuzzy detection on the images: first it loads and prepares the image segments, then find the local peaks, and performs

the fuzzy peak detection and fuzzy edge detection. The output is the coordinates of the pixels with high scores.

The components of the batch file are: FPD_INIT to define the variables; FPD to perform the fuzzy peak detection; FED to perform fuzzy edge detection; GRAD to perform gradient extraction; ISPEAK to test the local peaks; and THIN4 to thin the gradient images.

B.1.3 Analysis and comparison

The output of the detection program, in the form of a vector containing the location of the pixels and their respective score is processed by ROC01 and ROC02 to perform the ROC analysis. For comparison, the script NISHI is used to perform difference-image processing.

B.2 Program Listings

B.2.1 Openids3.m

```
function openids3(fname, xsize, ysize, nameroot, drive)
%OPENIDS3      Opens an IDS file
%       It opens the whole image and saves it in 500x500 chunks
%       Syntax: OPENIDS3(fname, xsize, ysize, nameroot)
%       eg. openids('aklcc.ids', 2702, 6160, 'aklc')
%       xsize and ysize are the image's dimension as given in the corresponding
%       .ics file.

tic;
idsfile= fname;
if nargin < 4 | nargin > 5
  disp('openids3(''fname'', xsize, ysize, ''nameroot'', ''drive'')')
  return;
end
if nargin < 5,
   drive = 'd';
```

```
end
eval(['!mkdir ' drive ':\' nameroot]);
J = ceil(xsize/500); %number of segments sideway
I = ceil(ysize/500); % ---------------  downwards
buff= zeros(500,xsize);

out = [];
file = fopen(idsfile, 'r');
for row = 1:ysize,
  [im_buff, im_size]= fread(file, xsize, 'uint16'); % read a line

  y = rem(im_buff,256);
  x = (im_buff-y)/256;

  lbuf = (y*256+x)'; % a line of pixels in raw format

  %save line buffer
  lname = sprintf('line%04d',row);
  eval([lname '=lbuf;']);

  if rem(row,500) == 0 | row == ysize, %process the strips into matrices
    iSeg = ceil(row / 500);
    iSize = 500;
    if row==ysize, iSize = rem(ysize,500); end
    for jSeg = 1:J, %create empty matrices
      jSize = 500;
      if jSeg==J, jSize =  rem(xsize,500); end
      sname = sprintf('%s%02d%02d', nameroot, iSeg, jSeg);
      eval([sname '=zeros(' num2str(iSize) ',' num2str(jSize) ');']);
      for i = 1:iSize, %xfer strips into matrices
        str = [nameroot ...
          sprintf('%02d%02d(%d,:)= ', iSeg, jSeg, i) ...
          sprintf('line%04d(%d:%d);', i+500*(iSeg-1), ...
            1+500*(jSeg-1), min(jSeg*500, xsize))];
        eval(str)
      end
      seg = sprintf('%s%02d%02d',nameroot,iSeg,jSeg);
      eval(['save ' drive ':\' nameroot '\' seg ' ' seg])
    end
      clear line*
      eval(['clear ' nameroot '*'])
  end

  if rem(row,250) == 0
    fprintf(1,'%4d ',row)
  end
  if rem(row,2000) == 0
    fprintf(1,'\n')
  end
end
fprintf(1,'\nfinish\n')
fclose(file);
fprintf(1,[showtime(toc) '\n'])
```

B.2.2 Opentru2.m

```
function [sig]= opentru2(fname, xsize, ysize, foutname)
%OPENTRU       Reads truth file in .ids format.
%       Returns location of signals in xxxx_mc and xxxx_cr files.
%       Syntax: opentru1(fname, xsize, ysize [, foutname])

if nargin < 3 | nargin > 4,
  disp('Insufficient or incorrect argument')
  return;
end
fout= 0;
if nargin == 4,
```

```
  fout= fopen(foutname, 'w');
  if fout == -1,
    disp('Unable to open file for output')
    return;
  else
    disp(sprintf('fid for fout=%d',fout))
  end
end
disp(sprintf('fid for fout=%d',fout))

idsfile= fname;
len= size(fname, 2);
if strcmp(fname(len-5:len-4),'cr') | strcmp(fname(len-5:len-4),'CR')
  iscr= 1;  ismc= 0;
  disp('it''s a cr file')
elseif (strcmp(fname(len-5:len-4),'mc'))
  disp('it''s an mc file')
  ismc= 1; iscr= 0;
else
  disp('Wrong filename or incompatible format.')
  return;
end

t0= clock; tnow= t0;
signal= [];
more off;
disp(sprintf('Start %s', showtime(etime(clock, t0))));

filein= fopen(idsfile, 'r');
if filein == -1,
  disp('Can''t open filein; exit')
  fclose all;
  return;
end
disp(sprintf('Filein opened, fid=%d %s',filein, showtime(etime(clock, t0))));

% Now read the words and break them into bits
for wordcount= 1: (xsize * ysize / 16),
  [bitte, sizen]= fread(filein, 1, 'uint16');       % read 16 bits (a word)
  if bitte ~= 0,
    if iscr == 1,
      if fout,
        fprintf(fout, '%d %d\n', ...
                rem(wordcount*16,xsize), floor(wordcount*16 / xsize)+1);
      else,
        signal= [signal; wordcount*16];
      end
    elseif ismc == 1,
      for i= 1:8,
        if bitte >= 2^(16-i),
          bitte= bitte - 2^(16-i);
          signal= [signal; (wordcount-1)*16 + i+8];
        end
      end
      for i= 9:16,
        if bitte >= 2^(16-i),
          bitte= bitte - 2^(16-i);
          signal= [signal; (wordcount-1)*16 + i-8];
        end
      end
    end
  end
end

fclose(filein);

if signal,
   sig= [rem(signal,xsize), floor(signal ./ xsize)+1,];
else
   sig = [];
end
```

```
disp(sprintf('Time: %s', showtime(etime(clock, t0))))

if (fout), fclose(fout); end
clear ans bitte filein fileout idsfile imels sizen
clear tnow wordcount xsize ysize
more on;

%end of file
```

B.2.3 Xyplot.m

```
function xyplot(data1, str1, data2, str2, data3, str3)
%XYPLOT Draws an x-y plot.
%       XYPLOT(M) draws the xy plot of M, where M is a matrix with
%       two columns; first column x data, second column y data.
%       XYPLOT(M,str) plots M with string specification str. Refer to
%       PLOT for descriptions of recognisable strings.
%       XYPLOT(M1,str1,M2,str2) plots M1 and M2.
%       A.P. Drijarkara 7 November 1997

if nargin == 1
   if size(data1,2) == 3,
      ax = axis; % = [xmin xmax ymin ymxax]
      for i = 1:size(data1,1),
         line(data1(i,1),data1(i,2),'marker','o','color','r',...
            'markersize',data1(i,3)*5+2);
         text('position',data1(i,1:2)+[1 1],...
            'string',num2str(data1(i,[2 1]),' %4d'),'fontsize',7);
         ax(1) = min(ax(1),data1(i,1));
         ax(2) = max(ax(2),data1(i,1));
         ax(3) = min(ax(3),data1(i,2));
         ax(4) = max(ax(4),data1(i,2));
         axis(ax);
      end
   else
      plot( data1(:,1), data1(:,2),'xk'), axis ij, axis image
   end
end

if nargin == 2
   if size(data1,2) == 3,
      for i = 1:size(data1,1),
         line(data1(i,1),data(i,2),'marker','o','color',str1,...
            'markersize',data1(i,3)*10);
      end
   else
      plot( data1(:,1), data1(:,2), str1), axis ij, axis image
   end
end

if nargin ==3
   [i,j]=ind2sub([str1,data2],data1);
   xyplot([i,j])
end

if nargin == 4
   plot( data1(:,1), data1(:,2), str1, ...
      data2(:,1), data2(:,2), str2),
   axis ij, axis image
end

if nargin == 6
   plot( data1(:,1), data1(:,2), str1, ...
      data2(:,1), data2(:,2), str2, ...
      data3(:,1), data3(:,2), str3),
   axis ij, axis image
end
```

```
if nargin > 6
       disp('Too many argument')
end

set(gca,'xtick',[0:500:6000],'ytick',[0:500:8000]);
grid on
```

B.2.4 Run01.m

```
% mamm is loaded by the script sizes.m
name = mamm(mg_idx).name ;    % name of case
xsize = mamm(mg_idx).xsize;   % x size of image
ysize = mamm(mg_idx).ysize;   % y size of image

% cases are stored in d:\<name>\ as 500x500 chunks, with filenames nameYYXX

% Prepare image segments, append the skirts from
% adjoining segments.

ysegmax = ceil(ysize/500); % No. of segment in y direction
xsegmax = ceil(xsize/500); % No. of segment in x direction
seg_count = 0;
swidth = 50;   % width of strips appended from adjoining segment
for iseg = 1:ysegmax,
  for jseg = 1:xsegmax,
    segname = [name num2str([iseg jseg],'%02d')];    % name of this segment
    eval(['load \' name '\' name num2str([iseg jseg],'%02d')]); % load this segment
    if iseg > 1, % load upper segment
      eval(['load \' name '\' name num2str([iseg-1 jseg],'%02d')]);
      eval(['segU = ' name num2str([iseg-1 jseg],'%02d') '(500-swidth+1:500,:);'])
      eval(['seg_count = seg_count + 1;'])
      eval(['clear '  name num2str([iseg-1 jseg],'%02d')]);
      if jseg > 1, % load upper-left segment
        eval(['load \' name '\' name num2str([iseg-1 jseg-1],'%02d')]);
        eval(['segUL = ' name num2str([iseg-1 jseg-1],'%02d') ';'])
        eval(['segUL = segUL(500-swidth+1:500,500-swidth+1:500);'])
        eval(['clear '   name num2str([iseg-1 jseg-1],'%02d')]);
      end
      if jseg < xsegmax, % load upper-right segment
        eval(['load \' name '\' name num2str([iseg-1 jseg+1],'%02d')]);
        eval(['segUR = ' name num2str([iseg-1 jseg+1],'%02d') ';'])
        eval('segUR = segUR(500-swidth+1:500,1:min(swidth,size(segUR,2)));')
        eval(['clear '   name num2str([iseg-1 jseg+1],'%02d')]);
      end
    end
    if iseg < ysegmax, % load lower/bottom segment
      eval(['load \' name '\' name num2str([iseg+1 jseg],'%02d')]);
      eval(['segB = ' name num2str([iseg+1 jseg],'%02d') ';'])
      eval(['segB = segB(1:min(swidth,size(segB,1)),:);'])
      eval(['seg_count = seg_count + 4;'])
      eval(['clear '  name num2str([iseg+1 jseg],'%02d')]);
      if jseg > 1, % load lower-left segment
        eval(['load \' name '\' name num2str([iseg+1 jseg-1],'%02d')]);
        eval(['segBL = ' name num2str([iseg+1 jseg-1],'%02d') ';'])
        eval(['segBL = segBL(1:min(swidth,size(segBL,1)),500-swidth+1:500);'])
        eval(['clear '   name num2str([iseg+1 jseg-1],'%02d')]);
      end
      if jseg < xsegmax, % load lower-right segment
        eval(['load \' name '\' name num2str([iseg+1 jseg+1],'%02d')]);
        eval(['segBR = ' name num2str([iseg+1 jseg+1],'%02d') ';'])
        eval(['segBR = ',...
          'segBR(1:min(swidth,size(segBR,1)),1:min(swidth,size(segBR,2)));'])
        eval(['clear '   name num2str([iseg+1 jseg+1],'%02d')]);
      end
    end
    if jseg > 1, % load left segment
```

```
      eval(['load \' name '\' name num2str([iseg jseg-1],'%02d')]);
      eval(['segL = ' name num2str([iseg jseg-1],'%02d') '(:,500-swidth+1:500);'])
      eval(['seg_count = seg_count + 2;'])
      eval(['clear '  name num2str([iseg jseg-1],'%02d')]);
    end
    if jseg < xsegmax, % load right segment
      eval(['load \' name '\' name num2str([iseg jseg+1],'%02d')]);
      eval(['segR = ' name num2str([iseg jseg+1],'%02d') ';'])
      eval(['segR = segR(:,1:min(swidth,size(segR,2)));'])
      eval(['seg_count = seg_count + 8;'])
      eval(['clear '  name num2str([iseg jseg+1],'%02d')]);
    end
    eval(['seg = ' segname ';'])
    eval(['clear ' segname])
    switch(seg_count),
    case 15
      eval(['seg = [segUL,segU,segUR;segL,seg,segR;segBL,segB,segBR];'])
      clear segUL segU segUR segL segR segBL segB segBR
    case 7
      eval(['seg = [segUL,segU;segL,seg;segBL,segB];'])
      clear segUL segU segL segBL segB
    case 11
      eval(['seg = [segUL,segU,segUR;segL,seg,segR];'])
      clear segUL segU segUR segL segR
    case 13
      eval(['seg = [segU,segUR;seg,segR;segB,segBR];'])
      clear segU segUR segR segB segBR
    case 14
      eval(['seg = [segL,seg,segR;segBL,segB,segBR];'])
      clear segL segR segBL segB segBR
    case 3
      eval(['seg = [segUL,segU;segL,seg];'])
      clear segUL segU segL
    case 5
      eval(['seg = [segU;seg;segB];'])
      clear segU segB
    case 9
      eval(['seg = [segU,segUR;seg,segR];'])
      clear segU segUR segR
    case 6
      eval(['seg = [segL,seg;segBL,segB];'])
      clear segL segBL segB
    case 10
      eval(['seg = [segL,seg,segR];'])
      clear segL segR
    case 12
      eval(['seg = [seg,segR;segB,segBR];'])
      clear segR segB segBR
    case 1
      eval(['seg = [segU;seg];'])
    case 2
      eval(['seg = [segL,seg];'])
    case 4
      eval(['seg = [seg;segB];'])
    case 8
      eval(['seg = [seg,segR];'])
    end
  end
end

clear ysegmax xsegmax xsize ysize name
```

B.2.5 Run03.m

```
kimage = 1;
% mamm is loaded by the script sizes.m
```

```
if ~exist('mamm'), sizes, end
name = mamm(kimage).name ;                    % name of image
xsize = mamm(kimage).xsize;                   % x size of image
ysize = mamm(kimage).ysize;                   % y size of image

ysegmax = ceil(ysize/500);                    % No. of segment in y direction
xsegmax = ceil(xsize/500);                    % No. of segment in x direction

fpd_init;                                     % initialise constants for fpd
if exist('peak')==1, clear peak, end
for iseg = 1:ysegmax,
  for jseg = 1:xsegmax,
    run01;              % segment is appended with strip from adjacent segments
    findpeak1           % find local peak; location in peak.i, peak.j
    ridge = thin4(grad(a));
    out = zeros(size(a));

    for k = 1:length(peak)
      i = peak(k).i; j = peak(k).j;
      peak(k).iabs = i - ioff + (iseg - 1) * 500;
      peak(k).jabs = j - joff + (jseg - 1) * 500;
      %now apply fpd; fpd returns m1, m2, m3 as fields of peak(k).
      fpd
      %and then edge detector; return peak(k).edge
      if peak(k).m1 > 0.2,
        fed
      else
        peak(k).edge = 0;
      end
    end
  end
end
```

B.2.6 Fpd_init.m

```
% Script for defining constants and coefficients for fpd
global mcwidth bright dark mNear mFar near
mcwidth = 10;                                 % half-width of mc
lpeaks = 5;                                   % half-width for local peaks
localArea = 50;

%bright =      [0.7; 0.9];
%dark =        [0.7; 0.9];
%near =        [2; 6];

[y,x] = meshgrid(-mcwidth:mcwidth, -mcwidth:mcwidth);
mNear = sqrt(y.^2 + x.^2);
mNear = zmf(mNear,near);
mFar = 1-mNear;
mFarSum = sum(mFar(:));
mNearSum = sum(mNear(:));

[y,x] = meshgrid(-localArea:localArea,-localArea:localArea);
dxy = sqrt(y.^2 + x.^2);
dxy(51,51) = inf;
```

B.2.7 Fpd.m

```
% Script to perform peak detection. Run with run03.m
% Inputs:
%              a: image segment
%              i,j: index of the candidate pixel in a
%              minA: the minimum in the area surrounding this object
```

```
%              mcwidth, bright, dark, mNear, mFar

% normalise ROI
minA = min(a(:));
g = rampmf(a(i-mcwidth:i+mcwidth, j-mcwidth:j+mcwidth),[minA a(i,j)]);

mBright = smf(g,bright);
mDark = zmf(g,dark);
mBN = min(mBright,mNear);
mDF = min(mDark,mFar);
mPeak = max(mBN, mDF);
peak(k).m1 = mean(mPeak(:));
peak(k).m2 = sum(mBN(:)) / mNearSum;
peak(k).m3 = sum(mDF(:)) / mFarSum;

out(i,j) = peak(k).m1;
```

B.2.8 Fed.m

```
% Script for detecting edge around peak object
g = ridge(i-localArea:i+localArea, j-localArea:j+localArea);
g = fedget2(g);
g = g ./ dxy;
t = 2;
g(51-t:51+t,51-t:51+t) = 0;
peak(k).edge = sum(g(:));
```

B.2.9 Grad.m

```
function G = grad(A0)
%GRAD   Gradient image
%       G = grad(M)
%       Returns the highest gradient among four directions around each pixel.
%       A.P. Drijarkara 21 Oct 1998

gFh= [0 0  0;...
      1 0 -1;...
      0 0  0]/2;
gFv=gFh';
gFd= [1 0  0;...
      0 0  0;...
      0 0 -1] / 2/sqrt(2);
gFu=flipud(gFd);
A1=abs(filter2(gFh,A0));
A2=abs(filter2(gFv,A0));
A3=abs(filter2(gFd,A0));
A4=abs(filter2(gFu,A0));
G =max(cat(3,A1,A2,A3,A4),[],3);
Gmin = min(G(:));
G(1:end,1) = Gmin;
G(1:end,end) = Gmin;
G(1,1:end) = Gmin;
G(end,1:end) = Gmin;
```

B.2.10 Ispeak.m

```
function out=ispeak(X)
%ISPEAK Check if matrix X has a peak in the center
```

```
%       X must be square and size must be odd number
%       Returns 1 if yes, 0 otherwise
%       A.P. Drijarkara 14 Aug 1998
dim = size(X,1);
if dim ~= size(X,2),
   disp('Matrix is not square')
   out = -1;
   return
end
if rem(dim,2) ~= 1,
   disp('Matrix size is not odd number')
   out = -1;
   return;
end
centre = (dim+1) / 2;
Xcentre = X(centre,centre);
if max(X(:)) > Xcentre %if another pixel is brighter, out.
   out = 0;
   return;
else
   if length(find(X==max(X(:)))) < 4 %if no more than 3 pixels == Xcentre
      out = 1;
   else
      out = 0;
   end
end
```

B.2.11 Thin4.m

```
function A1 = thin(A)
%THIN4  Thin greyscale image based on 2x2 window
%       AP Drijarkara 1 December 1998

A1 = A;
for i = 1:size(A,1)-1
   for j = 1:size(A,2)-1
      if A(i,j),
         Amin = A(i,j); Ai = i; Aj = j;
         if A(i,j+1) < Amin, Ai = i; Aj = j+1; Amin = A(i,j+1); end
         if A(i+1,j) < Amin, Ai = i+1; Aj = j; Amin = A(i+1,j); end
         if A(i+1,j+1) < Amin, Ai = i+1; Aj = j+1; end
         A1(Ai,Aj) = 0;
      end
   end
end
```

B.2.12 Roc01.m

```
%ROC    Perform ROC analysis of testdata.
% Data has the form: [mg_idx cr_idx yy xx mc_type fpeak fedge]
%                      1      2    3  4    5       6     7
%       mg_idx: the numeric code for mammogram image
%       mc_type: 1 = non-MC, 2 = maybe MC, 3 = MC, 4 = MC from truth file
%       AP Drijarkara 7 July 1999

true_mc = 4;
false_mc = 1;

P_data = data(find(data(:,5)==true_mc),:);
N_data = data(find(data(:,5)==false_mc),:);
data = [N_data; P_data];

% per-MC evaluation
```

```
P_mc = size(P_data,1);
N_mc = size(N_data,1);

max_th = ceil(max(data(:,7)));
fe_th = [0:max_th/200:max_th];
TPR_mc = []; FPR_mc = []; TPR_cl = []; FPR_cl = [];
TP_cl_lim = 1;
FP_cl_lim = 2;

for k = 1:length(fe_th)
  TP_mc = (data(:,5) >= true_mc) & (data(:,7) >= fe_th(k)) ;
  FP_mc = (data(:,5) < true_mc) & (data(:,7) >= fe_th(k));
  TPR_mc(k) = sum(TP_mc) / P_mc;
  FPR_mc(k) = sum(FP_mc) / N_mc;
  roc02;
end
plot(FPR_mc,TPR_mc,'.-r')
A_z=polyarea([1 FPR_mc ],[0 TPR_mc ])
title([method ' ' num2str(A_z,'%.2f')])
return
plot(FPR_cl,TPR_cl,'.-r')
polyarea([max(FPR_cl) FPR_cl]/max(FPR_cl),[0 TPR_cl])
plot(FPR_mg,TPR_mg,'.-r')
polyarea([1 1 FPR_mg ],[0 1 TPR_mg ])
```

B.2.13 Roc02.m

```
% Count FP clusters and TP clusters
TP_cl = 0; P_cl = 0; FP_cl = 0;
TP_mg = 0; P_mg = 0; FP_mg = 0; N_mg = 0;
cl_count = [];          %[mg_idx, P_cl, TP_cl, FP_cl]
mg_count = [];

sd =[1 3 5 14 16 25 33 34 39 40 51 52 64 65 84 88 89 90 91 92 94 95 96 97 ...
    98 119 120 121 122 151];
mg_idx_set = sd;

for mg_idx = mg_idx_set,                    % iterate on each image in set
  mg_data = data(find(data(:,1)==mg_idx),:); % find all data for this image
  TP_cl = 0; FP_cl = 0;
  P_cl = sum(unique(mg_data(:,2)) ~= 0);    % number of cluster in this image
  if P_cl > 0                               % if cluster exist, image is positive
    P_mg = P_mg + 1;
  else
    N_mg = N_mg + 1;
  end
  th_data = mg_data(find(mg_data(:,7) > fe_th(k)),:); % thresholded data
  if ~isempty(th_data),
    for cr_idx = unique(th_data(:,2))'      % for each cluster in this image:
      cr_data = th_data(find(th_data(:,2)==cr_idx),:);% data for this cluster
      if cr_idx == 0,                %find false cluster by measuring cart. dist.
        fcr_idx = -1;
        for i = 1:size(cr_data,1)-1,
          for j = i+1:size(cr_data,1),
            if cartdist(cr_data(i,3:4),cr_data(j,3:4)) < 100
              if cr_data(i,2) == 0 & cr_data(j,2) == 0, % if both are zero
                cr_data([i j],2) = fcr_idx;
                fcr_idx = fcr_idx - 1;
              end
              if cr_data(i,2) ~= 0 | cr_data(j,2) ~= 0, % if any is non-zero
                if cr_data(i,2) ~= 0 & cr_data(j,2) ~= 0, % if both are non-zero
                  old_idx = max(cr_data([i j],2));   %
                  new_idx = min(cr_data([i j],2));   % index for both
                  cr_data(find(cr_data(:,2)==old_idx),2) = new_idx;
                else
                  cr_data([i j],2) = min(cr_data([i,j],2));
                end
```

```
            end
          end
        end
      end
      x = cr_data(:,2); % cluster indices of the mc
      FP_cl = sum(sum(repmat(x,1,length(unique(x))) == ...
        repmat(unique(x'),length(x),1)) > FP_cl_lim & unique(x')~=0);
      fpcl_data = cr_data;
    else
      % count how many mc detected in this cluster; if > TP_cl_lim then TP_cl++.
      TP_cl = TP_cl + (size(cr_data,1) > TP_cl_lim);
    end
  end
  else % threshold is higher than any signal
    FP_cl = 0; TP_cl = 0;
  end
  % compile result for all images for this operating point:
  cl_count = [cl_count;mg_idx P_cl TP_cl FP_cl];
end
% now cl_count contains result for all images but only for one op. point
TPR_cl(k) = sum(cl_count(:,3)) / sum(cl_count(:,2));
FPR_cl(k) = sum(cl_count(:,4)) / length(mg_idx_set);
TPR_mg(k) = sum(cl_count(:,2) & cl_count(:,3)) / P_mg;
if N_mg,
  FPR_mg(k) = sum(~cl_count(:,2) & cl_count(:,4)) / N_mg;
else
  FPR_mg(k) = 0;
end
fappend('status.txt',[num2str(k) ' ' datestr(now) '\n']);
```

B.2.14 Nishi.m

```
function out = nishi(im);
%NISHI  Perform analysis as described by Nishikawa 95

w1 = 9; w2 = 31; w3 = (w2+1) / 2;
x1 = (w2-w1)/2; x2 = x1 + 1; x4 = w2 - x1 + 1; x3 = x4 -1;

enhf = zeros(w2); enhf(x2:x3,x2:x3) = 1; enhf = enhf / sum(enhf(:));
supf = ones(w2);  supf(x2-1:x3+1,x2-1:x3+1) = 0; supf = supf / sum(supf(:));
comf = enhf-supf; % combined suppresion and enhancement filter
out = filter2(comf,im);
out([1:w3,end-w3+1:end],:) = 0;
out(:,[1:w3,end-w3+1:end]) = 0;

ims = sort(out(:));
thrs = ims(round(length(ims)*.98));

out = out .* (out >= thrs);
return
```